

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/348713082>

# Determining the Working Behaviour of Hydraulic System Using Support Vector Machine

Chapter · January 2021

DOI: 10.1007/978-981-15-8025-3\_74

CITATIONS

0

READS

11

2 authors:



Jatin Prakash

Indian Institute of Technology Indore

6 PUBLICATIONS 11 CITATIONS

SEE PROFILE



Pavan K Kankar

Indian Institute of Technology Indore

101 PUBLICATIONS 1,851 CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:



Development of Induction-Conduction Based Material Deposition system for Metal Additive Manufacturing [View project](#)



Development of Adaptive Double Sided Incremental Forming Process for Dieless Manufacturing [View project](#)

# Determining the Working Behaviour of Hydraulic System Using Support Vector Machine



Jatin Prakash  and P. K. Kankar 

## 1 Introduction

The hydraulic systems are one of the most critical systems in the industries because of its better capabilities. The use of hydraulic systems varies from weight lifting using a single-stage hydraulic cylinder to the applications in heavy earth moving equipment [1]. A simple hydraulic system consists of the actuators and the hydraulic fluid flow circuit, whereas a complex system may consist of various working circuits like actuating circuit, cooling circuit, etc. Proper functioning of the cooling circuit is as desired as the functioning of the actuating circuit. Very few research progress has been made in the field of condition monitoring of the cooling circuit of the hydraulic system. Till date, attempts have been made to identify the leakage detection from hydraulic cylinders [2]. Besides, a neural network has been developed for the fault diagnosis of complex fluid power systems and pumping machinery [3, 4]. Observations of various literature suggest the implementation of different machine learning models for the fault categorization and diagnosis [5]. The process of gathering the information contained by the signals acquired in data-driven technique is called the feature extraction. To extract the useful information from the captured signal, it is a crucial step in the condition monitoring of mechanical component or machinery. Feature ranking techniques are espoused in order to indicate the importance of data. Information gain, Fisher score, Wilcoxon ranking, etc., have been used for feature selection for the condition monitoring of rotary bearing [6, 7]. These extracted features are fed to the system so as to monitor the working behaviour of

---

J. Prakash · P. K. Kankar (✉)

System Dynamics Lab, Discipline of Mechanical Engineering, Indian Institute of Technology Indore, Indore 453552, India

e-mail: [pkankar@iiti.ac.in](mailto:pkankar@iiti.ac.in)

J. Prakash

e-mail: [phd1901103007@iiti.ac.in](mailto:phd1901103007@iiti.ac.in)

© Springer Nature Singapore Pte Ltd. 2021

V. H. Saran and R. K. Misra (eds.), *Advances in Systems Engineering*,

Lecture Notes in Mechanical Engineering,

[https://doi.org/10.1007/978-981-15-8025-3\\_74](https://doi.org/10.1007/978-981-15-8025-3_74)

mechanical systems. In a recent study, a deep neural network model has been adopted to inspect the condition of the cooling circuit of a hydraulic system [8]. Model-based fault recognition technique and data-based fault detection technique are two of the most widely used fault detection techniques in mechanical machinery. Because of the complexity of the mathematical modelling, model-based fault detection cannot be frequently used, thus, later one is more preferred.

In this work, the support vector machine (SVM) has been employed to supervise the health condition of the hydraulic system. The cooling circuit of the hydraulic system has been kept under investigation using the pressure signals. Four different statistical features have been extracted from the raw pressure signal acquired. All the features have been fed to the classifier as input for the behaviour classification. The correlation and their dependency plot have been used to visualize the data features and relationship among them. To achieve better accuracy, two different methods, namely tenfold cross-validation and grid search cross-validation methods have been employed. The results show better accuracy is achieved in the grid search cross-validation method.

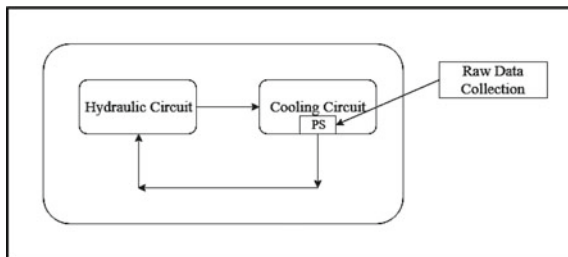
## 2 Methodology

### 2.1 Datasets

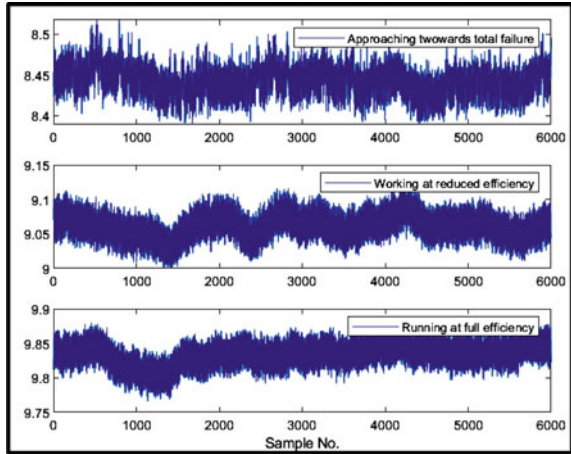
The predictive maintenance of the mechanical component or machinery can be performed using a historical dataset collected over time. Helwig et al. developed a test rig for the automated condition monitoring of the complex hydraulic system [9]. The whole circuit consists of various components. The data gathered for the pressure variation in the cooling circuit of the developed test rig has been employed in this study. The schematic view of the considered hydraulic system is shown in Fig. 1. 2205 instances of pressure variation data are gathered for 60 s at the sampling frequency of 100 Hz.

The flowing oil is entered in the cooling circuit after passing through a contamination filter and exits to the sink flowing out through a relief valve. The signals obtained from the pressure sensor belong to three dissimilar classes as mentioned

**Fig. 1** Block diagram showing the hydraulic and cooling circuit



**Fig. 2** Sampled signature of pressure variation [8]



by Prakash and Kankar [8]. Figure 2 shows the captured pressure signal with three diverse working conditions.

## 2.2 Feature Extraction

The very initial process to abstract out the useful information from the raw signals involves feature extraction from the signal. As such, there is no mathematical formulation to determine the optimal number of features. Thus, the selection of features to be extracted is of much significance. The features carry useful data about the raw data, but not all features are relevant. In [8], four optimal features have fed into the deep neural network and very good accuracy has been reported. So, in this study, four relevant features explained in brief below have been extracted.

- i. **Root Mean Square (RMS):** It is the square root of the arithmetic mean value of the squares of the signal readings. It is also known as the quadratic mean.

$$\text{RMS} = \sqrt{\frac{1}{n}(x_1^2 + x_2^2 + \cdots + x_n^2)} \quad (1)$$

- ii. **Standard Deviation:** It is static that measures the distribution of dataset with respect to its mean. It is the square root of the variance.

$$\text{Standard Deviation} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n - 1}} \quad (2)$$

- iii. **Kurtosis:** Kurtosis is derived from the Greek word “*Kurtos*” and it infers how heavily the tails of distribution differ from the tails of a normal distribution.

$$\text{Kurtosis} = \frac{\sum_{i=1}^n \frac{(x_i - \bar{x})}{n}}{\sigma^4} \quad (3)$$

- iv. **Skewness:** Skewness is the distortion or asymmetry in a symmetrical bell-shaped curve, mathematically known as the normal distribution. If the curve is shifted towards either end, then it is said to be skewed.

$$\text{Skewness} = \frac{\sum_{i=1}^n (x_i - \bar{x})^3}{n} \quad (4)$$

These four extracted features constitute the feature matrix. The sample feature values used to train the SVM have been mentioned in Table 1. The feature matrix is segregated into the ratio of 4:1 for the training and the testing purpose. Once the training of the classifier is finished, the testing data is fed to it.  $k$ -fold cross-validation method is employed to eliminate the biasness accompanied by the grid search method. The grid search method has been explained in Sect. 4.

In this work, tenfold cross-validation (for  $k = 10$ ) is used to visualize how the model performs when a new dataset is fed to it, whereas grid search is a way to choose the best among the family of models, parameterized using a grid of parameters.

**Table 1** Sample input feature values for SVM

	RMS	Standard deviation	Kurtosis	Skewness
Feature amplitude	9.185596	0.576802	1.541731	0.144612
	9.185666	0.577158	1.541128	0.142409
	9.185822	0.57733	1.541655	0.141214
	9.186052	0.57726	1.542304	0.140638
	9.186017	0.576846	1.5426	0.142243
	9.185874	0.576423	1.54277	0.143977
	9.185496	0.576332	1.542723	0.145696
	9.174669	0.573275	1.540543	0.151498
	9.174173	0.573551	1.540006	0.153071
	9.173866	0.573793	1.539785	0.153478
	9.173932	0.573795	1.540218	0.153755
	9.175028	0.572664	1.541951	0.15128
	9.174916	0.572707	1.541823	0.150424
	9.187213	0.574811	1.541986	0.148613
	9.187459	0.5747	1.541223	0.148226
	9.187597	0.574699	1.541951	0.147909
	9.187061	0.575101	1.541723	0.148787
	9.187267	0.575085	1.541632	0.14829
	9.187387	0.574991	1.542063	0.148439

**Fig. 3** Correlation matrix for extracted features

RMS	1.000	0.701	0.480	-0.516
ST DEV	0.701	1.000	-0.076	0.151
KURTOSIS	0.480	-0.076	1.000	-0.722
SKEWNESS	-0.516	0.151	-0.722	1.000
	RMS	ST DEV	KURTOSIS	SKEWNESS

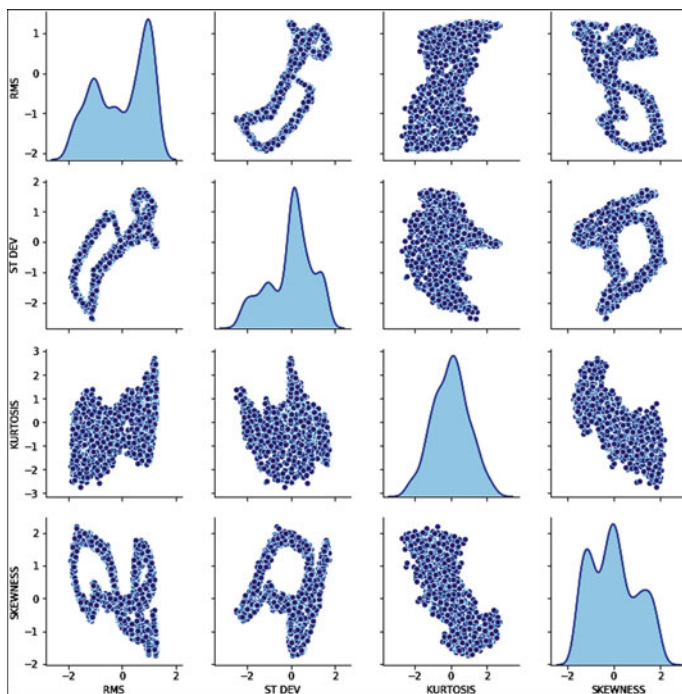
### 2.3 Correlation Coefficient

The correlation matrix is a table that shows the correlation coefficient between two variables. It is used to summarize a large amount of data to visualize the patterns and their mutual dependency. In the proposed manuscript, the correlation factor ( $r$ ) has been calculated for the four different extracted features with respect to each other and themselves also. It is obvious that the data depends on itself fully, so ' $r$ ' for itself is equal to unity and represented in diagonal elements. For two given variables, suppose  $x$  and  $y$ , the correlation coefficient can be calculated using Eq. 1. The correlation matrix of the extracted features is shown in Fig. 3. Higher values in the block show that these particular two properties are highly correlated.

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[n \sum x^2 - (\sum x)^2][n \sum y^2 - (\sum y)^2]}} \quad (5)$$

### 2.4 Dependency Plot

Figure 4 shows the plots among the four different features extracted from pressure signals. It can be treated as the data visualization technique. It shows the scattering of feature matrix data with respect to itself as well as with respect to each other feature. The histogram is obtained when distribution of feature vector is obtained with respect to itself. The histograms as shown in Fig. 4 can be generally classified



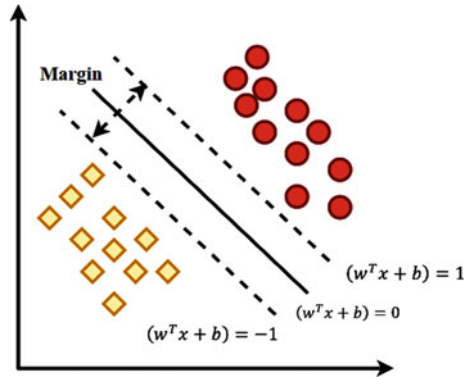
**Fig. 4** Dependency plot for four extracted features

into two classes, i.e. unimodal and bimodal histogram. Unimodal histogram infers the spread of data about a single value, whereas the bimodal histogram concludes the spread of data about two peak values. Most of the histograms (except Kurtosis) can be inferred as a bimodal histogram. Rest other corresponding plots show the distribution of the data with respect to another one. The trend observed in dependency plot also well satisfies the interpretations drawn from the correlation matrix in Sect. 2.3.

### 3 Support Vector Machine (SVM)

SVM is categorized as a supervised machine learning process proved efficient for the classification or regression problems. It exercises an extremely complex data transformation to define an optimal hyperplane between the possible categories of outputs using various kernels. SVM attempts to define a boundary between the two different categories in such a way that the marginal distance between two different classes is maximum, so that the generalization error is minimum. The adjacent data points used to describe the margin are known as support vectors. For example, in a given  $N$  number of sample set  $\{(x_i, y_i)\}; i = 1 \text{ to } N$ . It is required to determine a separation plane among different possible planes that distinct input samples into their

**Fig. 5** Representation of hyperplane classifying two-class problem by SVM



classes by least generalization error. For two different classes of data, the data can be labelled as  $y_i = 1$  and  $y_i = +1$  associated with the respective class. Slack variables ( $\xi$ ) are considered non-negative for a non-separable type of data. The hyperplane defined by  $f(x) = 0$  separating the given data can be attained as a solution using mentioned optimization problem [5].

$$\text{Minimize } \frac{1}{2} \|w\|^2 + c \sum_{i=1}^N \xi_i.$$

$$\text{subjected to } \begin{cases} y_i (w^T x_i + b) \geq 1 - \xi_i \\ \xi_i \geq 0, i = 1, 2, 3, \dots, N \end{cases} \text{ where } c \text{ is the error penalty constant.}$$

Figure 5 shows the two classes of data separated by a hyperplane.

## 4 Grid Search Method

The grid search method is an approach or process of scanning the data to configure optimal parameters for an artificially intelligent model. It tunes the methodically build model for various parameters and evaluate it with each combination of algorithm parameters stated in a grid. In this study, four different values of error penalty constant ( $c$ ) in combination with two different kernels have been employed. Nine discrete values of 'gamma' have been also used in the grids. Table 2 mentions the combination of different SVM hyperparameters used in the grid search method.

The higher value of the penalty parameter ( $C$ ) tries to minimize the misclassification of the training data, whereas the lower value of penalty parameter attempts to maintain a smooth classification. At the same time, if the value of the penalty parameter is set to be very high to correctly classify all the data, there is always a higher chance of overfitting in the model. So, in this study, the maximum value of  $C$  has been kept at 1000.

Gamma ( $\gamma$ ) is the inverse of the standard deviation of the Gaussian function that controls the trade-off between error due to bias and variance in the SVM model. For large values of  $\gamma$ , support vector does not have much impact on the classification of data.



**Table 2** Combination of SVM hyperparameters for grid search

Kernel	$C$ (Penalty parameter constant)
Linear	1
	10
	100
	1000
Rbf	1
	10
	100
	1000
Gamma	0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9

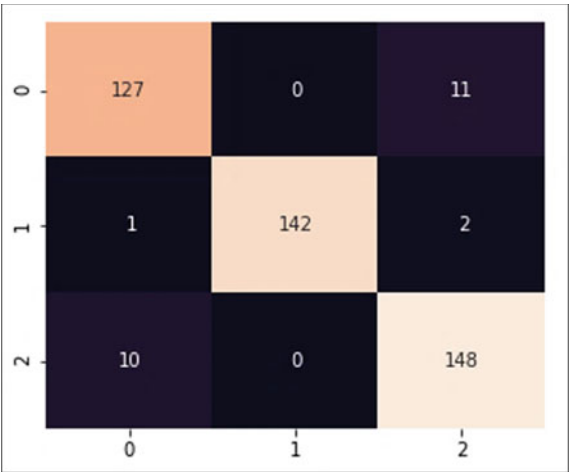
5 Results and Discussion

5.1 Confusion Matrix

A confusion matrix often called as an error matrix is generally used to designate the attainment of a classification model on a particular set of data for which the actual values are known. In the ongoing study, support vector classifier parameters are set as penalty parameter,  $C = 1.0$ , with ‘*rbf*’ kernel and the value of  $\gamma$  to be auto-deprecated. This classifier is used for the test data (20% of the total data, 441 in number) and the accuracy achieved is 95.91%. The corresponding confusion matrix is publicized in Fig. 6.

Further, to enhance the accuracy of the SVM classifier, tuning of the hyperparameters is required. In this work, tenfold cross-validation followed by grid search method has been performed. The cross-validation employing grid search finds out

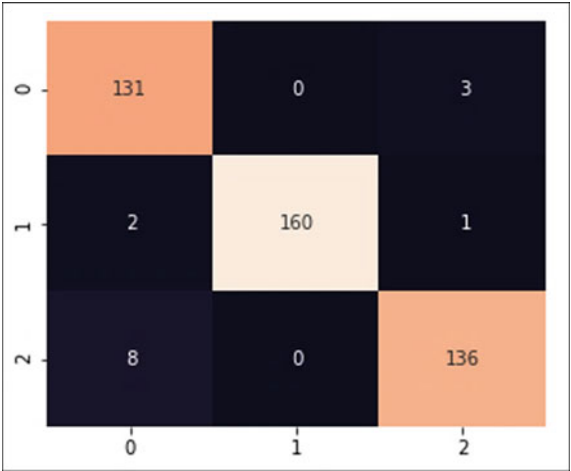
**Fig. 6** Confusion matrix for support vector classifier



**Table 3** Combination of the best hyperparameters according to grid search method

Hyperparameter	Value
Penalty parameter ( <i>C</i> )	1000
Gamma	0.6
Kernel	Rbf

**Fig. 7** Confusion matrix with the best parameters

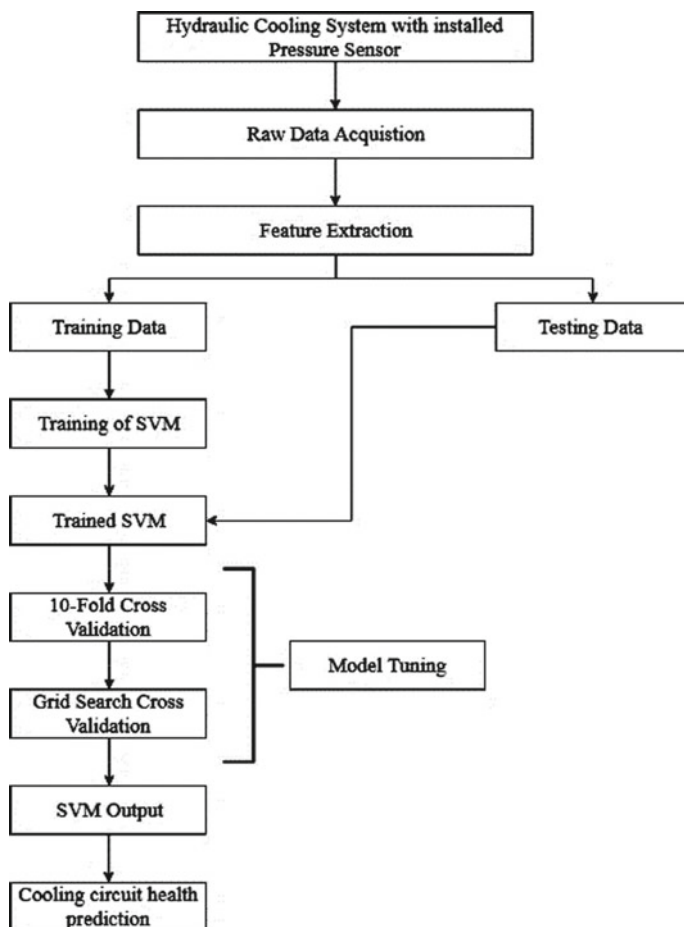


the best parameters for the better accuracy of the classifier by applying the various parameters to the classifier in one grid. The best parameters for this data in the support vector machine as determined by using grid search method are mentioned in Table 3. The combination of these parameters yields the accuracy of 98.80%. The corresponding confusion matrix is shown in Fig. 7. The flowchart of the methodology adapted is shown in Fig. 8.

6 Conclusions

In this study, the following conclusions and observations can be drawn:

- Working behaviour of the hydraulic cooling circuit has been determined using pressure signature employing support vector machine (SVM).
- The correlation matrix and the dependency plots have been shown in a good argument with each other.
- The grid search method can successfully be employed to determine the best hyperparameters combination.
- The accuracy has been boosted to 98.80% from 95.91% employing the best parameters as determined by grid search method.



**Fig. 8** Flowchart of cooling circuit condition monitoring

- The best accuracy of the developed model by Prakash and Kankar [8] using feature ranking technique is 99.54% which is not much less than that achieved in this study. Thus, it can also be concluded that feature ranking plays an important role in boosting the accuracy of the machine learning models.

## References

1. Prakash J, Gupta SK, Kankar PK (2020) An analytical approach to evaluate the maximum load carrying capacity for pin-mounted telescopic hydraulic cylinder. Proc Inst Mech Eng Part C J Mech Eng Sci, 0954406220916524

2. Yao J, Yang G, Ma D (2014) Internal leakage fault detection and tolerant control of single-rod hydraulic actuators. *Math Probl Eng*
3. Ramdén T, Krus P, Palmberg JO (1995) Fault diagnosis of complex fluid power systems using neural networks. In: *Proceedings of the fourth Scandinavian international conference on fluid power*, Tampere, Finland, pp 26–29
4. Ilott PW, Griffiths AJ (1997) Fault diagnosis of pumping machinery using artificial neural networks. *Proc Inst Mech Eng Part E J Process Mech Eng* 211(3):185–194
5. Vakharia V, Gupta VK, Kankar PK (2015) Ball bearing fault diagnosis using supervised and unsupervised machine learning methods. *Int J Acoust Vibr* 20(4):244–250
6. Upadhyay R, Manglick A, Reddy DK, Padhy PK, Kankar PK (2015) Channel optimization and nonlinear feature extraction for Electroencephalogram signals classification. *Comput Electr Eng* 45:222–234
7. Vakharia V, Gupta VK, Kankar PK (2016) Bearing fault diagnosis using feature ranking methods and fault identification algorithms. *Procedia Eng* 144:343–350
8. Prakash J, Kankar PK (2019) Health prediction of hydraulic cooling circuit using deep neural network with ensemble feature ranking technique. *Measurement*, 107225
9. Helwig N, Pignanelli E, Schütze A (2015) Condition monitoring of a complex hydraulic system using multivariate statistics. In: *2015 IEEE international instrumentation and measurement technology conference (I2MTC) proceedings*. IEEE, pp 210–215