JMPM: Jurnal Material dan Proses Manufaktur
Vol. 6, No. 2, p8-15, Desember 2022
DOI : https://doi.org/10.18196/jmpm.v6i2.16185

Vibration-Based Discriminant Analysis for Pipeline Leaks Detection

Berli Paripurna Kamiel[a], Indra Rukmana[b]
[a]Mechanical Engineering Department, Universitas Muhammadiyah Yogyakarta, Yogyakarta, Indonesia.
e-mail: berlikamiel@umy.ac.id

Kata kunci:

Akselerometer; *Discriminant Analysis*; Pipa, Getaran; *Machine learning*

ABSTRAK

Jalur pipa berguna memindahkan cairan dari satu tempat ke tempat lainnya. Masalah utama yang sering terjadi pada jalur pipa adalah kebocoran yang mengakibatkan kerugian produksi dan finansial. Pentingnya mendeteksi kebocoran jalur pipa membuat industri mencari metode deteksi yang efektif untuk menghindari kerugian yang lebih besar. Beberapa penelitian sebelumnya membuktikan bahwa metode berbasis getaran berhasil mendeteksi kebocoran pada jalur pipa. Namun demikian metode berbasis getaran yang digunakan tersebut rumit dan membutuhkan spesialis untuk menerjemahkan hasilnya. Penelitian ini mengusulkan sebuah metode deteksi berbasis machine learning yang mampu mengklasifikasi kondisi pipa secara langsung tanpa bantuan seorang spesialis. Metode yang diusulkan adalah *discriminant analysis* berbasis getaran; sebuah algoritma *machine learning* yang mengenali kondisi jalur pipa dari pola getarannya alih-alih dari spektrumnya. Metode yang diusulkan diuji pada sebuah *rig* uji yang terdiri dari jalur pipa *loop* tertutup yang dilengkapi sebuah segmen uji pipa bocor. Sinyal getaran diambil menggunakan sebuah akselerometer yang diletakkan pada segmen uji pipa bocor. Data getaran domain waktu diekstrak menggunakan parameter statistik yang bertujuan mengungkap informasi terkait kondisi pipa. Data getaran yang terkumpul dibagi menjadi dua kelompok yaitu data latih dan data uji. Model (*classifier*) berbasis *discriminant analysis* dilatih mengenali pola getaran jalur pipa menggunakan data-latih dan kemudian diuji kemampuannya mengklasifikasi data baru menggunakan data-uji. Terdapat empat ukuran kebocoran yang digunakan pada penelitian ini, kecil, sedang dan besar. Sedangkan kondisi normal (tidak bocor) digunakan sebagai benchmarking. Hasil penelitian menunjukkan bahwa metode yang diusulkan efektif mengklasifikasi empat kondisi pipa dengan akurasi mencapai 95%.

***Keyword*:**

Accelerometer; Discriminant Analysis; Pipe, Vibration; Machine learning

ABSTRACT

Pipelines are useful for transporting liquids from one place to another. The main problem that often occurs in pipelines is leakage which results in production and financial losses. The importance of detecting pipeline leaks makes the industries look for effective detection methods to avoid bigger losses. Several previous studies have proven that the vibration-based method is successful in detecting leaks in pipelines. However, the vibration-based method used in the previous study is relatively complicated and requires specialists to interpret the results. This study proposes a machine learning-based detection method that can classify pipe conditions directly without the help of a specialist. The proposed method is vibration-based discriminant analysis; a machine learning algorithm that recognizes pipeline conditions from their vibration pattern instead of spectrum. The proposed method was tested on a test rig consisting of a closed-loop pipeline equipped with a leak-pipe test segment. The vibration signal is taken using an accelerometer placed on the leak-pipe test segment. Time domain vibration data is extracted using several statistical parameters which aims to reveal information related to pipe conditions. The vibration data collected is divided into two groups, namely training-data and testing-data. The discriminant analysis model is trained to recognize the vibration pattern of the pipeline using training-data and then tested using testing-data. There are four leak sizes introduced in this study, small, medium, and large. Meanwhile, normal condition (no leaks) is used as benchmarking. The study shows that the proposed method is effective in classifying four pipe conditions with the accuracy up to 95%.

1. INTRODUCTION

Pipelines are the main system for moving liquid fluids from one place to another that are widely used in many companies and industries. The quality and quantity of water distribution is highly dependent on a reliable and efficient pipe network system [1]. However, in operation the pipeline network may fail which results in a decrease in the flow of water transferred. Pipeline damage may be in the form of pipe leaks along the distribution line. Several contributing factors include improper installation, corrosion, natural disasters or human error [2].

The main problem in diagnosing pipe damage is how to detect leaks and determine their size. Conventionally the location of a pipe leak is known visually by observing the formation of a puddle around the leak location. This method is inefficient and takes a long time to thoroughly inspect the pipeline. Therefore, we need a detection method that is effective, accurate and relatively easy to use by field operators.

Many studies have been conducted to obtain pipe leak detection using various methods, one of which is a method based on vibration signals. Yazdekhasti [3] managed to detect the location and size of pipe leaks using vibration spectra. Marmarokopos [4] proposed Fast Fourier Transform to detect leaky pipelines, similarly [5] proved that the vibration spectrum was successful in detecting leaks, the location and size of leaks in a pipeline..

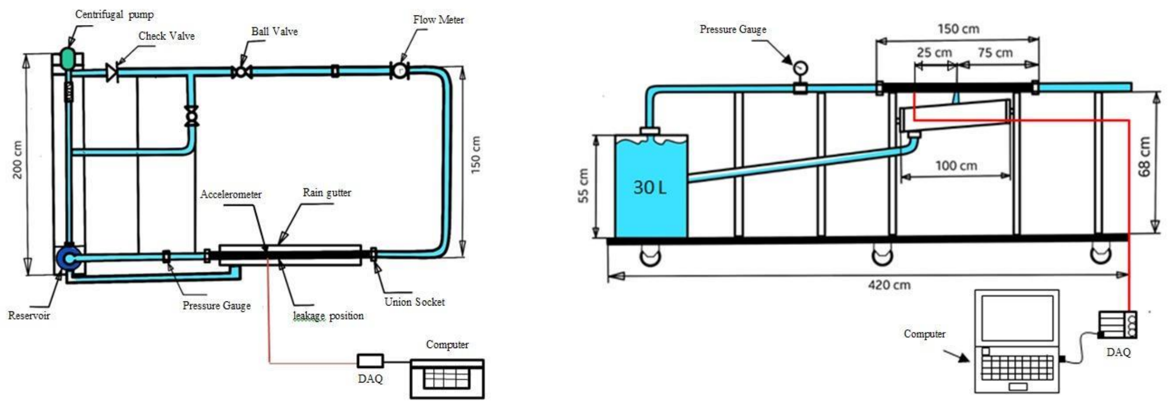
Almost all of the pipe leak detection proposals are based on the vibration spectrum[6], however, this method is relatively complicated and requires a specialist to translate and analyze the vibration spectrum. Another approach is based on machine learning. According to [7] machine learning is part of artificial intelligence that can recognize data patterns for prediction or classification purposes. There are many machine learning algorithms, one of which is discriminant analysis which can classify linear and non-linear data [8].

Several studies have proven that linear discriminant analysis (LDA) succeeded in classifying linear data and quadratic discriminant analysis (QDA) succeeded in classifying non-linear data. Jakovljevic [9] combined LDA and principal component analysis (PCA) to classify rotor damage in an induction motor. Likewise the study by [10] used LDA exclusively to detect rotor damage. Haddad [11] proposed the LDA method in combination with motor current signature analysis (MCSA) to detect damage to the motor magnet. Saputra [12] combined QDA and wavelet to classify misalignment of two shafts in induction motors.

Previous studies show that LDA and QDA are effective in classifying damage to induction motor components, however, there is no method based on discriminant analysis to classify or detect leaks in pipelines so there is a research gap that needs to be carried out immediately. This study proposes a leak detection method in water pipelines using discriminant analysis by utilizing vibration signals on the pipe walls. The aim of this research is to produce a detection method that is effective and easy to use by regular operators in the field.

2. METHOD

The test rig is a closed loop pipe network consisting of a series of 1 inch diameter PVC pipes, 1 HP centrifugal pump, a pressure gauge, a check valve, a ball valve, and a flow meter. The test pipe segments were made of 1-inch galvanized pipe 150 cm long, drilled to simulate a leak. There are four sets of test pipe segments used, they are normal pipe, small leaks (hole diameter 2 mm), medium leaks (hole diameter 5 mm) and large leaks (hole diameter 8 mm). The leak location is in the middle of the test pipe segment as shown in Figure 1. The data acquisition system consists of one Deltatron type 4507 accelerometer from Bruel & Kjaer with a sensitivity of 100.1 mV/g placed at a distance of 25 cm downstream of the leak location, one unit the NI 9234 data acquisition module from National Instrument and a computer equipped with data acquisition and signal processing application software.



(a) (b)

Figure 1. Test rig for simulating pipe leaks (a) top view and (b) front view

The accelerometer sensor records vibration signals with a sampling rate of 25600 Hz with a recording duration of 20 seconds. The number of records taken for each pipe condition is 600. Fourteen statistical parameters, namely mean, RMS, standard deviation, shape factor, kurtosis, skewness, crest factor, clearance factor, variance, peak value, impulse factor, square root amplitude and maximum were extracted from the raw vibration signal. Ninety percent of the data is used as training data while the rest is used as test data. The flow chart of the proposed pipe leak detection method can be seen in Figure 2. The study proposes two types of discriminant analysis algorithms, namely linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA). Parameter selection was carried out using principal component analysis (PCA) and binomial theorem.

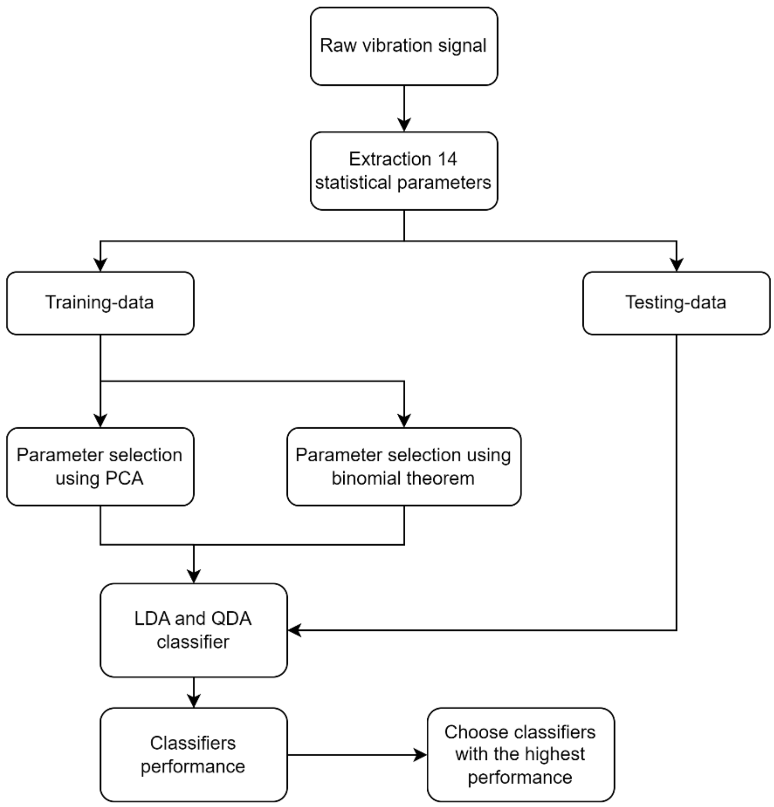


Figure 2. Flowchart of the proposed method

3. Result And Discussion

The four pipe leak conditions in the test section, namely normal, small leaks, medium leaks, and large leaks can be seen in Figure 3, while the corresponding time domain is shown in Figure 4.

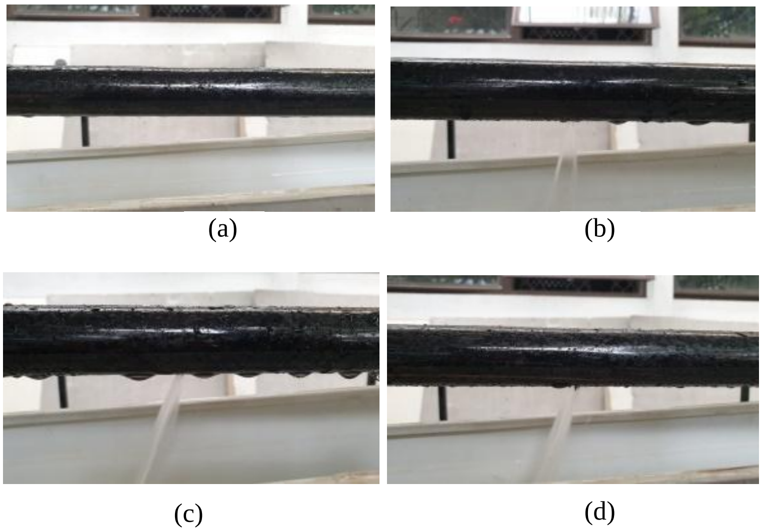


Figure 3. Test pipe, a) normal, b) small leak, c) medium leak, d) large leak

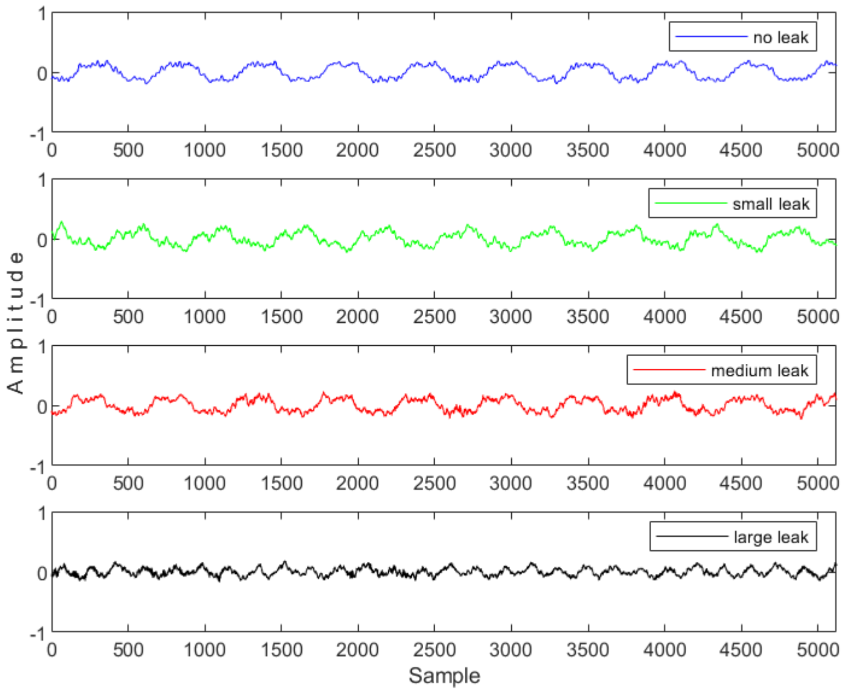


Figure 4. Time domain of pipe vibration

Figure 4. shows that there is no observed difference in vibration amplitude between the pipe conditions. For all pipe conditions, the vibration amplitude was at the level of 0.2 units with a slight increase in frequency for large leaky pipe vibrations, but in the other three conditions the observed vibration amplitudes tended to be the same. In terms of the shape/pattern of the time domain graph, there is also no significant difference between the four pipe conditions, so it is difficult to determine the condition of the pipe only from the vibration characteristics using time domain.

To reveal the hidden vibration characteristics, fourteen statistical parameters are extracted from the time domain and the results are plotted in the form of a matrix as shown in Figure 5 (black dot is normal, green dot is small leak, red dot is medium leak and blue dot is large leak). The matrix in Figure 5 shows that there is a tendency for the time domain statistical values of the four pipe conditions to converge into four clusters. namely clusters of normal pipe, small leaks, medium leaks, and large leaks. This proves that the value of the vibrational time domain statistical parameters has unique characteristics so that it has the potential to be used as the parameter of a classifier.

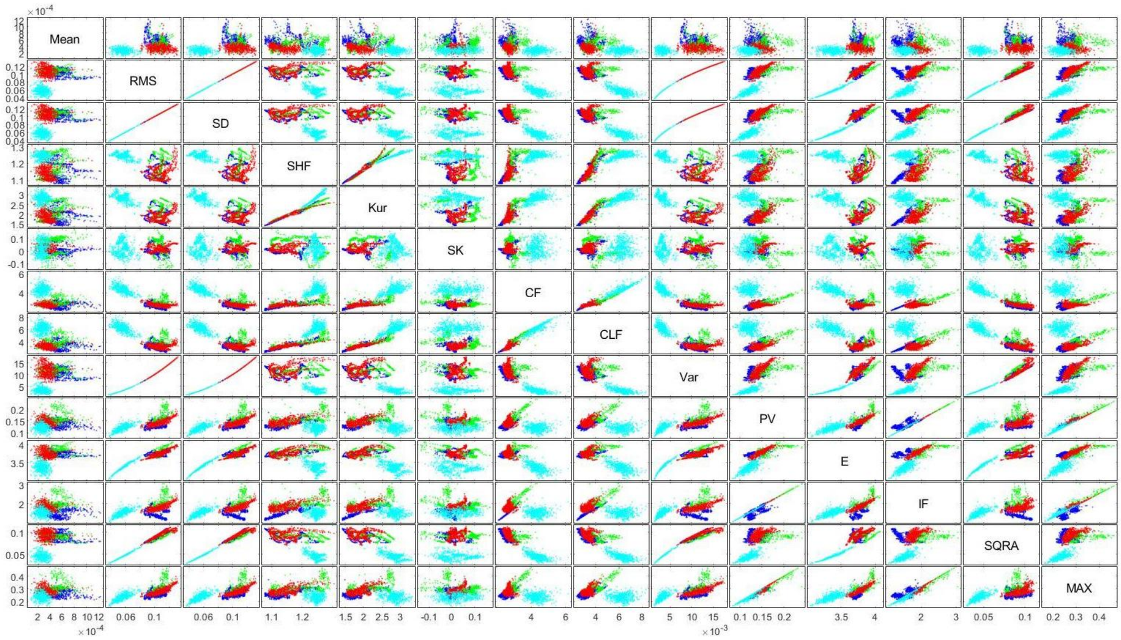


Figure 5. Matrix of statistical parameters

LDA is a classifier that works by maximizing the ratio of the between-class variance to the within-class variance. Fourteen statistical parameters are used as input for the LDA classifier by using 2160 training data and producing a classifier model. The model obtained was then tested using 240 test data. The results of classification accuracy can be seen in Table 1.

Table 1. Accuracy of LDA classifier

Input	Accuracy	
	Training-data	Testing-data
14 Statistical parameters	87,6%	90,4%

Table 2. Confusion matrix of training-data LDA classifier

		Normal	Small leak	Medium leak	Large leak
True class	Normal	57	6	0	0
	Small leak	1	42	12	0
	Medium leak	0	4	57	0
	Large leak	0	0	0	61
		Predicted class			

Classification results using the LDA model based on 14 statistical parameters achieved an accuracy of 87.6% for training data and 90.4% for test data. These results indicate that the LDA model is general in nature and does not experience overfitting. Table 2 presents the confusion matrix which states the classification performance of the LDA model for test data. It can be seen that there is a misclassification in the Normal, small leak and medium leak classes, but there is no misclassification in the large leak class. These results indicate that the LDA model is not sensitive to small changes in vibration characteristics due to small leaks and is only sensitive to vibrations originating from large leaks. However, increasing the level of accuracy can be done by selecting statistical parameters and eliminating redundant parameters. The selection process was tried for the first time using the trial-and-error method based on the binomial theorem. This method seeks to find a set of parameters that contribute the most to the LDA model building stage and provide higher accuracy. The selection results revealed that 11 statistical parameters namely RMS, standard deviation, shape factor, kurtosis, skewness, crest factor, variance, entropy, impulse factor, square root amplitude, and maximum produced an LDA model with higher accuracy. Classification accuracy for training data and testing data reached 87.8% and 91.3%, respectively.

The second selection method used is PCA-based selection. The first stage of the selection process using PCA is to normalize the data using the z-score which aims to eliminate the influence of various units in each statistical parameter. Then the data from the original domain is transformed into the principal component (PC) domain into 14 PCs which are sorted from PC1 which has the highest variance to PC14 which has the lowest variance. The variance value is directly proportional to the quality of information owned by a PC.

The transformation into the principal component domain shows that the four PCs with the highest variance (PC1 to PC4) contain 96.5% meaningful information, while the rest is information in the form of interference or noise. The accumulated variance values of the four PCs can be seen in the Pareto diagram in Figure 6.

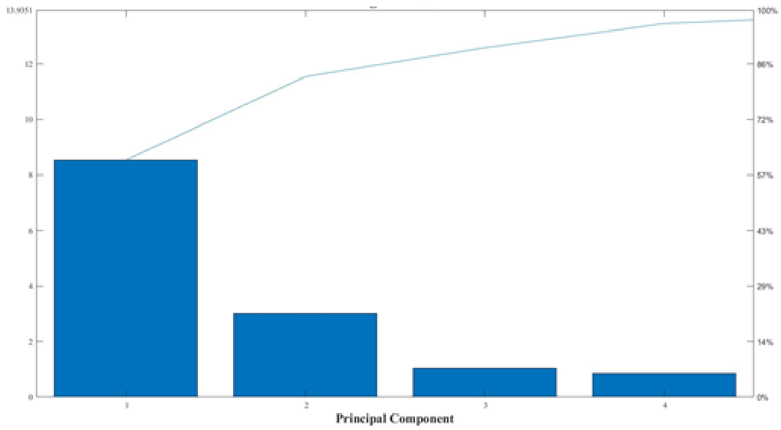


Figure 6. Pareto Diagram

The four selected PCs were then used to form the LDA-PCA model. The classification accuracy of the training-data and test-data of the LDA-PCA model is shown in Table 3, while the test-data confusion matrix is presented in Table 4. The classification results obtained were lower than the accuracy obtained from the previous method. Likewise, the confusion matrix in Table 4 shows more misclassifications compared to the previous LDA model. This result means that the transformation into the principal component domain and dimension reduction to 4 PCs do not provide the advantage of information accumulation. This is consistent with the results of the accumulated variance value of 4 PCs which amounted to only 96.5%. Statistically, this means that the 14 parameters in the original domain are not significantly related to each other, so the consequence is that the transformation into the PC domain is not successful in gathering the majority of information to the first few PCs.

Table 3. Accuracy of LDA-PCA classifier

Input	Akurasi	
	Training-data	Testing-data
4 *principal components*	85,1%	89,2%

Table 4. Confusion matrix testing-data LDA-PCA classifier

		Normal	Small leak	Medium leak	Large leak
True class	Normal	57	6	0	0
	Small leak	0	43	12	0
	Medium leak	0	8	53	0
	Large leak	0	0	0	61
		Predicted class			

Classification using the quadratic discriminant analysis (QDA) algorithm provides accuracy as presented in Table 5. It can be seen that QDA without parameter selection stages provides the highest accuracy compared to the two previous methods. The accuracy value for both training data and test data reaches 95% which is obtained using 13 input parameters at the modeling stage. One parameter, namely the impulse factor, was excluded from the group because it resulted in a singularity at the modeling stage. No overfitting symptoms were identified in this model, as evidenced by the accuracy value which did not differ in classification using training data and test data. The QDA confusion matrix in Table 6 shows misclassification in the normal class, small leak, and medium leak in a row of 4, 5 and 3 data while in the large leak class there is no classification error. The results obtained are in line with the results of the two previous models, where misclassification occurred in the normal class, small leak, and medium leak. This indicates that the pipe vibration signals under normal conditions, small leak and medium leak tend not to have strong characteristics so that it is relatively difficult to classify. The results of the large leak class classification for all models show high accuracy which means that the vibration of the large leak pipe is significantly different from the other 3 classes. In general, it can be concluded that all models have higher accuracy than previous studies [13, 14] so that the method proposed in this study is promising for application in the field.

The results obtained by QDA reveal that the data set has a non-linear tendency which causes the LDA to not provide high accuracy. Many misclassifications occur in the LDA model, even with the help of the PCA selection procedure.

Table 5. Accuracy of QDA classifier

Input	Akurasi	
	Training-data	Testing-data
13 Statistical parameters	95%	95%

Table 6 Confusion matrix of QDA

		Normal	Small leak	Medium leak	Large leak
True class	Normal	59	2	2	0
	Small leak	1	50	4	0
	Medium leak	2	1	58	0
	Large leak	0	0	0	61
		Predicted class			

4. CONCLUSION

In this study, discriminant analysis classification method and statistical parameter selection technique from vibration time domain are proposed to detect leaks in water pipelines. The conclusions from the research results can be expressed in two aspects, (1) the discriminant analysis algorithm is effective in classifying vibration signals in leaky pipes with an accuracy of up to 90.4% for test data. The increase in classification accuracy with the parameter selection method was successfully carried out using the binomial method with an accuracy of 91.3%, while PCA-based selection with four PCs was not effective in increasing the accuracy value with an accuracy value of 89.2%. (2) Classification using a non-linear algorithm, namely QDA, achieves the highest accuracy value, namely 95% for test data. This accuracy is obtained by the QDA model without using parameter selection. The better classification accuracy achieved by the QDA algorithm indicates that the pipe leakage vibration data set is non-linear.

REFERENCES

[1] Mudiarto, A., Supriyadi, and Sugiyanto. Pemodelan Fisik Untuk Monitoring Kebocoran Pipa Air Dengan Metode Geolistrik. *Unnes Physics Journal*, 2012; 1.

[2] Rahmat, R.F., I. Satria, B. Siregar, and R. Budiarto. Water Pipeline Monitoring and Leak Detection using Flow Liquid Meter Sensor. *IOP Conference Series: Materials Science and Engineering*, 2016; 190.

[3] Yazdekhasti, S., K.R. Piratla, S. Atamturktur, and A.A. Khan. Novel vibration-based technique for detecting water pipeline leakage. *Structure and Infrastructure Engineering*, 2017; 13: 731 - 742.

[4] Marmarokopos, K., D. Doukakis, G.N. Frantziskonis, and M. Avlonitis. Leak Detection in Plastic Water Supply Pipes with a High Signal-to-Noise Ratio Accelerometer. *Measurement and Control*, 2018; 51: 27 - 37.

[5] Okosun, F., P. Cahill, B. Hazra, and V. Pakrashi. Vibration-based leak detection and monitoring of water pipes using output-only piezoelectric sensors. *The European Physical Journal Special Topics*, 2019; 228(7): 1659-1675.

[6] Yazdekhasti, S., K. Piratla, S. Atamturktur, and A. Khan. Experimental evaluation of a vibration-based leak detection technique for water pipelines. *Structure and Infrastructure Engineering*, 2017; 14: 1-10.

[7] Sharma, P. and M. Kaur. Classification in Pattern Recognition: A Review. *International Journal of Research in Computer Science and Software Engineering*, 2013; 3(4): 298-306.

[8] Tharwat, A. Linear vs. quadratic discriminant analysis classifier: a tutorial. *Int. J. Appl. Pattern Recognit.*, 2016; 3: 145-180.

[9] Jakovljevic, B.B., Z. Kanovic, and Z.D. Jeličić. Induction motor broken bar detection using vibration signal analysis, principal component analysis and linear discriminant analysis. *2012 IEEE International Conference on Control Applications*, 2012: 1686-1690.

[10] Fernandez-Temprano, M., P.E. Gardel-Sotomayor, O. Duque-Perez, and D. Morinigo-Sotelo. *Broken bar condition monitoring of an induction motor under different supplies using a linear discriminant analysis*. in 2013 9th IEEE International Symposium on Diagnostics for Electric Machines, Power Electronics and Drives (SDEMPED). 2013.

[11] Haddad, R.Z. and E.G. Strangas. On the Accuracy of Fault Detection and Separation in Permanent Magnet Synchronous Machines Using MCSA/MVSA and LDA. *IEEE Transactions on Energy Conversion*, 2016; 31: 924-934.

[12] Saputra, P.P.S., F.D. Murdianto, R. Firmansyah, and K. Widarsono. Combination Of Quadratic Discriminant Analysis and Daubechis Wavelet For Classification Level Of Misalignment On Induction Motor. *2019 International Symposium on Electronics and Smart Devices (ISESD)*, 2019: 1-5.

[13] Tsitsifli, S., V. Kanakoudis, and I. Bakouros. *Pipe reliability assessment using discriminant analysis and classification: a case study from Mexico*. in Protection and Restoration of the Environment VIII. 2006. Chania, Crete, Greece

[14] Chi, Z., Y. Li, W. Wang, C. Xu, and R. Yuan. Detection of water pipeline leakage based on random forest. *Journal of Physics: Conference Series*, 2021; 1978(1): 012044.