



Combination of MFCC Feature Extraction and Support Vector Machine for Pipeline Leak Detection

Feby Anggraeini Rosyifah^{1,*}, Anindita Adikaputri Vinaya¹, and
Andhika Eko Prasetyo¹

¹ Department of Engineering Management, Internasional Semen Indonesia University,
Complex PT. Semen Indonesia, Veteran Street, Gresik 61122, Indonesia

Corresponding Author's email: feby.rosyifah@student.uisi.ac.id

Abstract. Pipes have an important role in industries that support fluid or gas distribution. Therefore, it is necessary to periodically monitor the condition of the pipe, especially if what is flowed can be flammable or dangerous if it comes out of the pipe. Leaks in the pipe need to be avoided. In the era of the Industrial Revolution 4.0, condition monitoring by utilizing sensors as data collectors that are processed and machine learning as a condition classifier is being developed. The sound produced by the pipe will be extracted using the Mel-Frequency Cepstral Coefficients (MFCC) feature which will be the input for machine learning. MFCC features are considered to have high sensitivity which is expected to result in high model performance. This research will compare the performance of the model used for classification in 2 (two) different environmental conditions, namely laboratory scale and workshop scale. The SVM-based machine learning model can predict the condition of leaks and no leaks as evidenced by the percentage of F1-score performance that reaches 95% at the laboratory scale and 90.95% at the workshop scale. the model has a high percentage of accuracy which is 89.74% at the laboratory scale and 92.17% at the workshop scale. The general objective of this research is to test the performance of the combination of MFCC feature extraction with the Support Vector Machine model in detecting pipe leaks in environments with different noise levels.

Keywords: Pipeline, Leak, MFCC, Machine Learning, SVM.

1 INTRODUCTION

Pipes are defined as tubular products with a function to distribute fluids and gases which have an important role in supporting the production system of a factory [1]. With this role, pipeline equipment, especially gas pipes, requires continuous monitoring to avoid leaks that can disrupt the production process. With the growth of service time, the pipes are gradually aging, or are corroded by various media and other damage factors [2]. One simple way to detect leaks is with sound, which has been done by several studies using sound to detect anomalous conditions that occur in a piece of equipment [3]. In the Industry 4.0 era, machine conditions can be detected based on

data or audio signals from microphones as sensors combined with machine learning [4].

The use of machine learning is highly beneficial for monitoring the condition of gas pipelines. This is because relying solely on human hearing for leak detection is considered inadequate, given the limitations of the human auditory range, which has a minimum threshold of 20 Hz and a maximum threshold of 20,000 Hz or 20 kHz [5]. With these limitations, the likelihood of operators or maintenance personnel detecting differences in the sound of a leaking pipeline is very low. This is further exacerbated by the variable locations of pipelines, which can be in places with low noise levels such as laboratories, or high noise levels such as workshops. Furthermore, if we rely solely on a detection system based on the sound spectrum generated by a leaking pipeline, it would require operators in the control room to be extremely attentive in monitoring and analyzing sound spectra quickly [3]. At times, during their work, operators may make human errors that result in an incorrect and hasty analysis of pipeline conditions by the sound of gas pipes. Human errors occur due to various factors, one of which is physiological factors caused by the monotony of their tasks, leading to reduced concentration [6]. Whenever failures occur, with a percentage of around 70-80%, humans become the main cause of equipment failures, incidents, and accidents [7]. Based on these two situations, the use of machine learning is advantageous and can reduce human errors as well as errors in detecting gas pipeline leaks.

Machine learning is one of the tools that can be used to classify the condition of a machine or equipment. Condition classification can be done based on audio signals extracted using Mel-Frequency Cepstral Coefficients (MFCC) features. Feature extraction and representation significantly impact the performance of machine learning models. Mel Frequency Cepstral Coefficient (MFCC) is designed for feature extraction of audio signals and is widely used in various fields [8]. In several studies that design a speech recognition system, MFCC features are widely used as machine learning input because only a few studies report good results when using time-domain features [8]. The MFCC feature has been widely used in previous studies and has good results on speech signals and rotating engine sounds [9]. Studies related to the use of the MFCC feature on stationary equipment such as pipes are still limited. In another study, MFCC was used as an extraction feature in language speech recognition with model performance during training reaching more than 80% [10].

In this study, MFCC will be used as input for machine learning with the Support Vector Machine (SVM) classification model. Comparison of model performance for tube leak detection is carried out on 2 (two) different noise scales, namely laboratory and workshop. The comparison is carried out to determine the percentage change in model performance if there are differences in environmental conditions. This research is expected to be a new reference related to pipe leak detection with a combination of MFCC extraction features and machine learning.

2 METHOD

2.1 Mel-Frequency Cepstral Coefficient (MFCC)

MFCC is most commonly used in speech recognition because MFCC can work well on inputs with a high level of correlation by removing information that is not needed during the extraction process [10]. MFCC is a method that can represent the power of sound that works like the human sense of hearing [9]. Because of this, MFCC is often used as a method for speech recognition. Applications of MFCC include speech recognition, speaker recognition, emotion recognition, bearing fault detection, and gear fault detection[8]. The sound frequency extracted using MFCC can be converted into a mel-scale using the following equation [9]:

$$\text{Mel}(f) = 2595 * \log_{10}(1 + f/700) \tag{1}$$

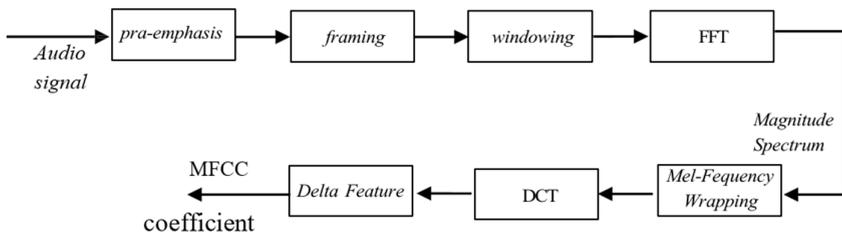


Fig. 1. The flow of MFCC Feature Extraction [8]

Based on Figure 1, The MFCC feature extraction process begins with the initial stages such as improving the quality of audio files to the Delta Feature stage which aims to obtain information from dynamic features which will then be added to the feature vector to produce a more accurate feature vector. In this stage, the delta coefficient will be generated as much as the cepstral coefficient generated by MFCC. The following are the stages in obtaining the MFCC features [8]:

- a. Pre-emphasis, this stage is the initial stage of acoustic signals from audio recordings in the form of datasets will be carried out a quality improvement process.
- b. Framing, at this stage the acoustic signal will be divided into several frames and overlapping frames will be carried out with the aim that the information available in the audio recording is not lost.
- c. Windowing, at this stage the acoustic signal will begin to be filtered to minimize distortion between frames which is done by multiplying between frames with the type of window used. The following is a formula for windowing:

$$w(k) = 0,54 - 0,46 \cos\left(\frac{2\pi nk}{K-1}\right) \tag{2}$$

Description:

N = number of samples

n = window index

K = number of frames

- d. Fast Fourier Transform (FFT), this stage will make changes to each frame from the time domain to the frequency domain. FFT is an algorithm that implements a fast version of the Discrete Fourier Transform (DCT).

$$f(n) = \sum_{k=0}^{N-1} y_k e^{-\frac{2\pi}{N}kn} \quad (3)$$

Description:

N = number of samples

n = window index

y = windowing result signal

- e. Mel-Frequency Wrapping, in this stage the energy size of the frequency wrapping in the acoustic signal will be known. Mel Frequency Wrapping is generally done using Filter banks, which have the aim of knowing the energy of certain frequency bands in acoustic signals. With the reception of acoustic signals in audio recordings for low frequencies (<1000 Hz) which are linear and high frequencies (>1000 Hz) which are logarithmic. The following is the formula used in the calculation of filter banks:

$$(f) = 1125 * \ln(1 + f/700) \quad (4)$$

Description:

f = linear frequency (Hz)

B(f) = mel-frequency scale

- f. Discrete Cosine Transform (DCT), is the final stage of the MFCC main process for feature extraction. The DCT stage uses the basic concept of correlating mel spectra to produce a good representation of spectral properties. The output of this stage is close to principle component analysis (PCA), which is a classical statistical method widely used for data analysis and compression. The following is the formula for calculating DCT:

$$C_s(n; m) = \sum_{k=0}^{N-1} a_k \cdot \log(fmel_k) \cos\left(\frac{\pi(2n+1)k}{2N}\right) \quad (5)$$

Description :

N = Number of samples

n = 0,1,2,3...N-1

fmel = Frequency mel

- g. Delta Feature, this stage aims to obtain information from dynamic features which will then be added to the feature vector to produce a more accurate feature vector. In this stage, the delta coefficient will be generated as much as the cepstral coefficient generated by MFCC. The formula of the delta feature is as follows:

$$d_t = \frac{\sum_{n=1}^N n(C_{t+n} - C_{t-n})}{2 \sum_{n=1}^N n^2} \quad (6)$$

Description :

2.2 Machine Learning

The trend of using modeling in data analysis is widely adopted by various companies to predict events [11]. Machine learning is the science of computer algorithms that are included in the artificial intelligence (AI) section [9]. One of the

methods is classification which produces new data groups from old data groups that are already labeled as known supervised learning techniques [12]. This method operates by training the model with multiple labeled datasets to generate new data as the output of the prediction [12]. One of the models that falls under supervised learning is the Support Vector Machine model, which can be used for detection and prediction based on historical data. In machine learning, data is divided into testing and training data sets. The testing data is used to train the model and comprises the bulk of the dataset. While the testing dataset is used to evaluate the performance of the model [13]. Model evaluation measurements using the evaluation metrics or confusion matrix described below.

Table 1. Confusion Matrix [13].

Actual Class	Prediction Class	
	Positive	Negative
Positive	True Positive (TP)	False Negative (FN)
Negative	False Positive (FP)	True Negative (TN)

Table 1 shows the confusion matrix, which is a comparison between the actual class or events that occurred in the field and the predicted class based on the patterns and characteristics of the input. The evaluation matrix has the appearance of a combination of predicted and actual values. The table of the confusion matrix contains an N x N matrix where the N value indicates the class label used in modeling [13]. There are several terms in the confusion matrix including true positive, true negative, false positive, and false negative.

$$accuracy = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \tag{7}$$

Accuracy is the ratio of the total correct predictions to the total subject values. The accuracy formula is shown in 7.

$$recall = \frac{TP}{TP+FN} \times 100\% \tag{8}$$

Recall, or referred to as sensitivity is the ratio of the ratio between True Positives (TP) and the amount of data that is actually positive. The smaller the False Negative (FN), the greater the recall. The formula of recall is shown in 8.

$$specificity = \frac{TN}{TN+FP} \times 100\% \tag{9}$$

Specificity is the True Negative (TN) ratio of all subjects who do not have a disability or condition. The specificity formula is shown in 9.

$$precision = \frac{TP}{TP+FP} \times 100\% \tag{10}$$

Precision, the ratio between True Positive (TP) and the number of data predicted to be positive, the smaller the FP, the greater the precision. The precision formula is shown in 10.

$$f1-score = 2 * \frac{precision*recall}{precision+recall} \times 100\% \tag{11}$$

F1-Score by definition is the harmonic mean of precision and recall. The resulting output is a range of 0 to 1, where 0 describes the bad model used and 1 describes the best model.

2.3 Support Vector Machine

The Support Vector Machine (SVM) is capable of solving two types of problems, namely classification and regression. Typically, the SVM method can be used for two or more problems, such as multi-class scenarios [14]. Modeling using the SVM algorithm requires data that can be linearly separable, allowing it to be divided into two classes with a hyperplane [9]. For data that is not linearly separable, a kernel function will be used, as it is considered more flexible and capable of handling nonlinear issues, as shown in Figure 3.

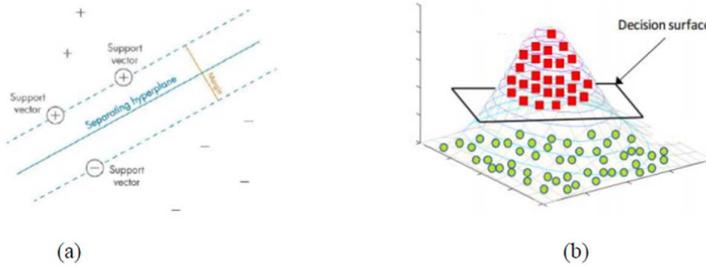


Fig. 2. (a) The Support Vector Machine classification standards and (b) kernel functions [15]

During the classification process, the Support Vector Machine has several types of classifiers and algorithms, including linear, polynomial, Gaussian, and sigmoid [16]. The research focuses on three types of classifiers, with explanations provided in Table 1. An example of using the SVM model for pipe leak detection is shown in the research of Sohaib & Kim (2019) entitled "Data-Driven Leakage Detection and Classification of a Boiler Tube" which conducted a leak test on a boiler tube analyzed using wavelet packet transform (WPT) analysis and using acoustic emission (AE) to capture signals from the boiler tube [17]. In the analysis, to improve the classification performance of leak identification, it is proposed to use the K-Nearest Neighbor (KNN) and Support Vector Machine (SVM) models which are able to produce an average percentage accuracy of 99.2% when testing with signals captured by AE [17]. This shows that the SVM model is an effective classification model used in leak testing for boiler tubes.

Table 2. Classifier of Support Vector Machine [16]

Classifier	Interpretability	Model Flexibility
Linear SVM (LSVM) 	Easy	Low Makes a simple linear separation between classes
Medium Gaussian SVM (MG SVM) 	Hard	Medium Medium distinctions, with kernel scale set to sqrt (P)
Optimizable SVM (OSVM)	Hard	Medium Using optimal hyperparameters

2.4 Dataset and Research Flowchart

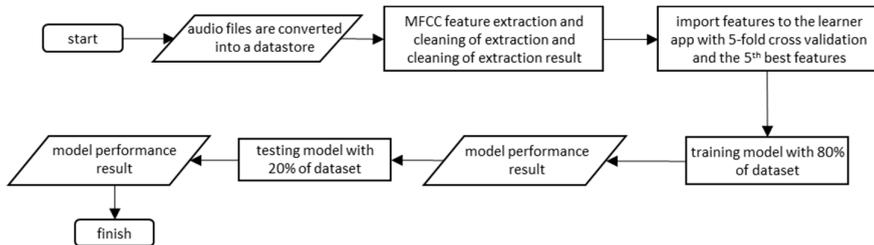


Fig. 3. Data processing scheme.

Figure 4 illustrates the data processing flow, starting from extraction to the output of performance values from the model used, which is the Support Vector Machine (SVM) during the training and testing processes. The data used in the research is audio file data published by the Fraunhofer Institute for Digital Media (IDMT), Industrial Media Applications (IMA) Group in the IDMT-ISA-COMPRESSED-AIR project conducted by Johnson, et al (2020) [4]. The audio file contains tube leakage data which is divided based on the noise scale, namely laboratory and workshop. In the laboratory scale tube leak audio file, there is no high noise so the pipe leak sound will be more audible. This research uses a frequency range of 22.05 kHz or 22050 Hz to get a high-accuracy model obtained from half the standard frequency sample rate of 44100 Hz which is slightly lower than the actual audio frequency sample rate of 48000 Hz, where the maximum frequency range of 44100 Hz is 22050 Hz [9]. The total dataset in the form of audio files used as training and testing data at each scale (laboratory and workshop) is 192 audio files with details of 80% of the dataset or 154 data used as training data and the remaining 20% or 38 data used as testing data.

The first step according to the flowchart is the audio files are imported into the datastore. The next step is to split data into 2 types, 80% of the total data will be used as training data and 20% of the remaining will become testing data. The training data will be used to train a model that uses most of the dataset so the model can determine the patterns and characteristics of the pipe leakage data. Next, the audio files are extracted to get the MFCC features and continue by selecting the 5 of 14 (X1, X2, X3, ..., X14, X15) best features based on the Maximum Relevancy Minimum Redundancy (MRMR) algorithm that will be used in classification. This is based on previous research by Berg-Hansen (2022), where MFCC features are used as a method for extracting audio files and the 5th best features are used to get high accuracy [9]. The code of extraction MFCC features is shown in Figure 5 using f_s (frequency sample) 22050 Hz or 22,05 kHz.

Extract Features - MFCC

```

fs = 22050; %22.05khz
adstall = tell(trainDatastore);
% melspectrum, barkSpectrum, erb5Spectrum
aFE = audioFeatureExtractor('SampleRate',22.05e3,'mfcc',true,'pitch',true, ...
    'Window',hamming(round(fs),'periodic'), ...
    'OverlapLength',fs*0.5);
specsTall = cellfun(@(x)extract(aFE,x),adstall,'UniformOutput',false);
specs = gather(specsTall);
  
```

Fig. 4. The Code of MFCC Extraction Feature.

The result of extraction is shown in Figure 6 which contains the 14 features from MFCC as predictors and 1 label column as the categorical prediction. The next step is to classification of the training data with the Support Vector Machine (SVM) models with Linear SVM, Medium Gaussian SVM, and Optimizable SVM classifiers. The classification training results are displayed in the form of a confusion matrix. The model obtained will be tested with other data which will also produce a confusion matrix that will be used for performance calculation.

Table 3. The Result of the MFCC Extraction Process

	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	X14	X15
1	-6.14E+14	2.89E-01	2.82E+14	1.40E+14	8.35E-01	5.34E-01	5.28E-01	4.46E-01	4.49E-01	4.49E-01	3.07E-01	3.43E-01	3.18E-01	3.95E+02	'Leak'
2	-5.09E+14	1.28E+14	2.45E+14	1.84E+14	4.79E-01	6.40E-01	3.92E-01	4.56E-01	4.52E-01	4.21E-01	3.07E-01	3.64E-01	2.28E-01	4.00E+02	'Leak'
3	-6.62E+14	2.81E+14	2.38E+14	1.53E+14	4.70E-01	6.43E-01	4.48E-01	3.34E-01	3.77E-01	2.20E-01	2.98E-01	2.40E-01	3.61E-01	4.00E+02	'Leak'
4	-5.62E+14	2.88E+14	2.32E+14	1.57E+14	8.96E-01	6.15E-01	4.07E-01	2.78E-01	2.93E-01	1.27E-01	2.42E-01	1.77E-01	3.42E-01	4.00E+02	'Leak'
5	-6.34E+14	1.63E+14	2.66E+14	1.63E+14	6.39E-01	9.27E-01	-8.06E-02	3.00E-01	6.21E-01	-1.79E-01	4.66E-01	1.65E-01	2.59E-01	3.94E+02	'Leak'
304	-6.74E+14	3.54E+14	1.98E+14	1.55E+14	2.79E-01	5.92E-01	3.65E-01	2.60E-01	3.82E-01	1.33E-01	4.39E-01	1.70E-01	2.31E-01	4.00E+02	'No Leak'
305	-5.72E+14	3.78E+14	1.88E+14	1.61E+14	1.22E-01	6.14E-01	3.27E-01	2.05E-01	2.90E-01	3.24E-02	3.51E-01	1.49E-01	2.24E-01	4.00E+02	'No Leak'
306	-6.72E+14	4.11E+14	2.15E+14	1.88E+14	1.06E-01	6.56E-01	2.77E-01	1.45E-01	4.31E-01	-1.18E-02	4.07E-01	1.58E-01	-1.16E-02	3.82E+02	'No Leak'
307	-5.98E+14	3.93E+14	2.20E+14	1.81E+14	-4.12E-02	7.97E-01	1.71E-01	1.26E-01	4.56E-01	-2.40E-01	5.92E-01	-2.07E-02	3.75E-02	3.66E+02	'No Leak'
308	-5.69E+14	3.85E+14	1.73E+14	1.62E+14	9.94E-02	7.52E-01	3.71E-01	2.83E-01	3.76E-01	1.69E-02	4.74E-01	1.50E-01	2.41E-01	4.00E+02	'No Leak'

3 RESULT AND DISCUSSION

The results of feature extraction are imported into the learner app for the classification process, with the requirement of using a 5-fold validation technique. The cross-validation technique is to estimate the performance of a model when applied to testing data, ensuring that predicted values do not differ significantly from the results obtained during model training [18]. The next step is to select three classifiers: Linear SVM, Medium Gaussian SVM, and Optimizable SVM, to assess the performance during training utilizing 80% of the entire dataset at each noise level scale.

3.1 Laboratory Scale

On a laboratory scale, the noise setting of the test site is minimal as it is in a laboratory room [4]. The best feature selection results (X2, X13, X11, X5, X4) are used as Support Vector Machine (SVM) input. The machine will learn from a dataset that contains the 5th best features from MFCC as predictors and 1 label column as pipeline condition. The results of machine learning from the training data provided can produce the confusion matrix in Figure 7.

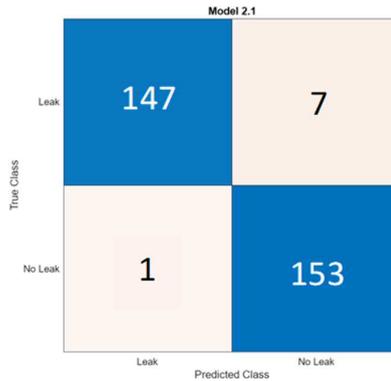


Fig. 5. Confusion Matrix of Training at Laboratory Scale.

The classification results shown in Figure 7 are examples of the confusion matrix generated after the training process. To understand the results, the vertical axis represents the labels of the actual tube conditions, while the horizontal axis represents the model's predicted outcomes. The differences in the model's predicted outcomes are influenced by the hyperparameters of the classifier used. Based on the results of the confusion matrix, 7 audio data files (pink color) were predicted by the model to have no leakage, while the actual pipeline had a leakage. On the other hand, 147 audio data files (blue color) were correctly predicted according to the actual condition of the pipe. In the next stage, the results from each classifier are exported to the workspace to enter the testing phase using a 20% dataset. The code for testing data is shown in Figure 8 using a sampling rate of 22,05 kHz.

```

fs = 22050; %22.05khz
adstall_test = tall(testDatastore);
% melspectrum, barkSpectrum, erbSpectrum
aFE =
audioFeatureExtractor('SampleRate',22.05e3,'mfcc',true,'pitch',true,
...
    "Window",hamming(round(fs),"periodic"), ...
    "OverlapLength",round(0.5*fs));
specsTall_test =
cellfun(@(x)extract(aFE,x),adstall_test,"UniformOutput",false);
specstest = gather(specsTall_test);

```

Fig. 6. Code for Testing data at Laboratory Scale.

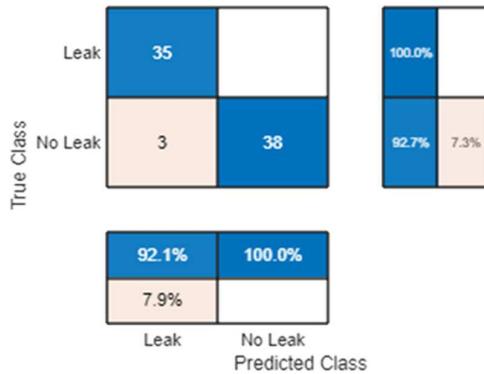


Fig. 7. Confusion Matrix of Testing at Laboratory Scale.

Figure 7 shows the display of the confusion matrix generated after testing. The result of confusion matrix testing is based on the data imported by the training confusion matrix to derive accurate testing values. From the image, it can be observed that there are 3 audio files (pink color) that have prediction errors. The model predicted leakage in the tube whereas, in reality, the tube is not leaking. On the other hand (blue color), there is consistency between the model's predicted outcomes and the actual conditions. A total of 38 and 35 audio files indicate that the model successfully predicted the tube's condition accurately. From the two confusion matrices above (training and testing), the next step taken is to calculate the performance values of the model for each classifier using formulas 7 to 11.

Table 4. The Result of the MFCC Extraction Process

Classifier	Categorical of Model Performance	Training	Testing	Average
Linear SVM (LSVM)	Accuracy	88,16%	90,00%	89,08%
	Recall	78,95%	90,00%	84,47%
	Specificity	97,37%	90,00%	93,68%
	Precision	96,77%	90,00%	93,39%
Medium Gaussian SVM (MGSVM)	F1-Score	97,07%	90,00%	93,54%
	Accuracy	88,16%	85,00%	86,58%
	Recall	78,95%	88,89%	83,92%
	Specificity	97,37%	81,82%	89,59%
	Precision	96,77%	80,00%	88,39%
Optimizable SVM (OSVM)	F1-Score	97,07%	80,90%	88,98%
	Accuracy	89,47%	90,00%	89,74%
	Recall	78,95%	90,00%	84,47%
	Specificity	100,00%	90,00%	95,00%
	Precision	100,00%	90,00%	95,00%
	F1-Score	100,00%	90,00%	95,00%

Table 4 presents the performance calculation results of the model based on 5 evaluation categories, including accuracy, recall, specificity, precision, and F1-score. Based on Table 1, it can be observed that there is a decrease in percentage from training to testing, as seen in the LSVM classifier, which has an F1-Score value of 97,07% in training, dropping to 90,00% in testing. This decrease is considered normal and does not indicate underfitting or overfitting, as the observed decrease is not substantial [19].

The decrease in percentage during testing indicates that the model can effectively classify conditions [13].

The assessment of the classifier's quality is based on the F1-score value, as well as the accuracy value, to gauge the classifier's ability to predict the dataset under two different conditions. From the performance calculation results of both training and testing, the average values will be taken as the final determinants of performance percentage for each classifier. For the Linear SVM classifier, the average F1-score stands at 93,54%, with an accuracy rate of 89,08%. Other classifiers, such as Medium Gaussian SVM, have average F1-scores and accuracy of 88,98% and 86,58% respectively. Furthermore, among the three classifiers used, the Optimizable SVM classifier has the highest F1-score value, which is 95,00%.

3.2 Workshop Scale

At the workshop scale, other noise is added from recorded sound as if it were in an actual workshop with a high noise level [9]. The data processing carried out to find the performance evaluation value uses the same method as the laboratory scale.

Table 5. Performance Percentage of Workshop Scale.

Classifier	Categorical of Model Performance	Training	Testing	Average
Linear SVM (LSVM)	Accuracy	90,26%	93,42%	91,84%
	Recall	88,31%	97,14%	92,73%
	Specificity	92,21%	90,24%	91,23%
	Precision	91,89%	89,47%	90,68%
	F1-Score	92,05%	89,86%	90,95%
Medium Gaussian SVM (MGSVM)	Accuracy	90,26%	93,42%	91,84%
	Recall	88,96%	100,00%	94,48%
	Specificity	91,56%	88,37%	89,97%
	Precision	91,33%	86,84%	89,09%
Optimizable SVM (OSVM)	F1-Score	91,45%	87,60%	89,52%
	Accuracy	90,91%	93,42%	92,17%
	Recall	90,26%	100,00%	95,13%
	Specificity	91,56%	88,37%	89,97%
	Precision	91,45%	86,84%	89,14%
	F1-Score	91,50%	87,60%	89,55%

Table 5 presents the performance calculation results of the model based on the generated confusion matrix values at the workshop scale. The Linear SVM classifier in the training and testing columns shows an increase in the accuracy percentage from 90,26% in training to 93,42% in testing. This can be indicative of positive performance, as it demonstrates the model's effective classification capability. However, this increase also suggests that the classifier might be using relatively simple hyperparameters, during testing, the model detects patterns in a straightforward and less complex manner [20].

To determine the quality of the classifier used, it can be assessed by the F1-score value which ranges from 0 to 1. If the classifier's percentage approaches 1 (100%), the model can be considered to predict the dataset effectively. Another assessment can be made using the accuracy value, indicating the classifier's ability to predict the dataset under two different conditions. From the performance calculation results of both

training and testing, the average values will be taken as the final determinants of performance percentage for each classifier. The Linear SVM classifier has an average F1-score value of 90,95%, indicating good model performance, and an accuracy value of 91,84%, demonstrating the model's accurate prediction of pipe conditions. Additionally, among the three classifiers used, the Linear SVM classifier has a higher F1-score value compared to the Medium Gaussian SVM, which has a percentage of 89,52%, and the Optimizable SVM with a percentage of 89,55%.

Based on the average values obtained for each category in both scales (laboratory and workshop), the next step involves comparing the performance percentages between the laboratory scale and the workshop scale. The primary assessment to determine the percentage decrease is based on the F1-score category, followed by the accuracy category, and finally, the specificity category to observe the percentage of normal conditions (No Leak) within the entire dataset.

Table 6. Comparison of Performance Between Laboratory Scale and Workshop Scale.

Classifier	Categorical of Model Performance	Laboratory	Workshop	Percentage Difference
Linear SVM (LSVM)	Accuracy	89,08%	91,84%	2,76%
	Recall	84,47%	92,73%	8,26%
	Specificity	93,68%	91,23%	2,45%
	Precision	93,39%	90,68%	2,71%
	F1-Score	93,54%	90,95%	2,59%
Medium Gaussian SVM (MGSVM)	Accuracy	86,58%	91,84%	5,26%
	Recall	83,92%	94,48%	10,56%
	Specificity	89,59%	89,97%	0,38%
	Precision	88,39%	89,09%	0,70%
	F1-Score	88,98%	89,52%	0,54%
Optimizable SVM (OSVM)	Accuracy	89,74%	92,17%	2,43%
	Recall	84,47%	95,13%	10,66%
	Specificity	95,00%	89,97%	5,03%
	Precision	95,00%	89,14%	5,86%
	F1-Score	95,00%	89,55%	5,45%

Based on Table 6, looking at the F1-score category in the Linear SVM classifier, there is a decrease in percentage from the laboratory scale, dropping from 93,54% to 90,95% in the workshop scale, with a difference of 2,59%. The Optimizable SVM classifier exhibits a relatively high difference, which is 5,45% for the F1-score category, decreasing from 95,00% on the laboratory scale to 89,55% on the workshop scale. The decrease observed is not excessively high and still falls within the percentage range of 80%, indicating that the model remains quite effective even when predicting under two different conditions.

On the other hand, despite the decrease in the F1-score percentage from the laboratory scale to the workshop scale, there is an increase in the accuracy category, as demonstrated by the Medium Gaussian SVM classifier. The accuracy shows a significant increase from 86,58% in the laboratory scale to 91,84% in the workshop scale. This accuracy improvement indicates that the model can predict the dataset correctly even in environments with high noise interference.

4 CONCLUSION

The average F1-score in the highest category obtained in the laboratory scale is 95% for the OSVM classifier, and the lowest is 88.98% for the MGSVM classifier, which can still be categorized as a model capable of detecting gas pipe leaks. Meanwhile, in the workshop scale, the highest average F1-score is 90.95% for the LSVM classifier, and the lowest is 89.52% for the MGSVM classifier. In other performance categories such as accuracy, the highest score for the laboratory scale is 89.74% for the OSVM classifier, and for the workshop scale, the highest percentage is 92.17% for the OSVM classifier. Based on these percentages, it can be demonstrated that the combination of MFCC and SVM can detect pipe leaks in two different environments with varying noise levels, namely the laboratory scale (low noise) and the workshop scale (high noise).

REFERENCES

1. A. Dey, "Types of Pipes: Classification of Pipes." Accessed: Aug. 12, 2023. [Online]. Available: https://whatispiping.com/types-of-pipes/?expand_article=1
2. Y. Zhao, F. Ma, N. Zhao, and J. Ma, "Study on Acoustic Detection Method of Gas Pipeline Leakage," in IOP Conference Series: Earth and Environmental Science, IOP Publishing Ltd, Sep. 2020. doi: 10.1088/1755-1315/558/2/022032.
3. M. S. Khan, T. Ashuri, M. Collins, G. Birkemeier, and E. Almohammedali, "An Acoustic-Based Approach for Condition Monitoring of Pipes," J Ark Acad Sci, vol. 71, Jan. 2017, doi: 10.54119/jaas.2017.7114.
4. D. Johnson, J. Kirner, S. Grollmisch, and J. Liebtrau, "Compressed Air Leakage Detection Using Acoustic Emissions with Neural Networks," in The 49th International Congress and Exposition on Noise Control Engineering (Inter-Noise), Seoul: Fraunhofer Institute for Digital Media (IDMT), Aug. 2020.
5. D. Purves et al., Neuroscience, 2nd Edition, 2nd ed. Sinauer Associates, 2001. Accessed: Oct. 25, 2023. [Online]. Available: <https://www.ncbi.nlm.nih.gov/books/NBK10799/>
6. W. Shi, F. Jiang, Q. Zheng, and J. Cui, "Analysis and control of human error," in Procedia Engineering, 2011, pp. 2126–2132. doi: 10.1016/j.proeng.2011.11.2415.
7. M. Nkosi, K. Gupta, and M. Mashinini, "Causes and Impact of Human Error in Maintenance of Mechanical Systems," MATEC Web of Conferences, vol. 312, p. 05001, 2020, doi: 10.1051/mateconf/202031205001.
8. Z. Kh. Abdur and A. K. Al-Talabani, "Mel Frequency Cepstral Coefficient and its Applications: A Review," IEEE Access, vol. 10, 2022, doi: 10.1109/ACCESS.2022.3223444.
9. S. Berg-Hansen, "An exploration of acoustic gas leak and anomaly detections suitability for autonomous inspections in industrial plants," 2022. Accessed: Feb. 23, 2023. [Online]. Available: <https://hdl.handle.net/11250/3024842>
10. W. Mustikarini, R. Hidayat, and A. Bejo, "Real-Time Indonesian Language Speech Recognition with MFCC Algorithms and Python-Based SVM," 2019.
11. G. G. Samatas, S. S. Moumgiakmas, and G. A. Papakostas, "Predictive Maintenance-Bridging Artificial Intelligence and IoT," May 2021. doi: <http://dx.doi.org/10.1109/AIIoT52608.2021.9454173>.

12. R. Thupae, B. Isong, N. Gasela, and A. M. Abu-Mahfouz, "Machine learning techniques for traffic identification and classification in SDWSN: A survey," in Proceedings: IECON 2018 - 44th Annual Conference of the IEEE Industrial Electronics Society, Institute of Electrical and Electronics Engineers Inc., Dec. 2018, pp. 4645–4650. doi: 10.1109/IECON.2018.8591178.
13. S. Amirneni, "Anomaly Detection in Highly Imbalanced Dataset," 2019.
14. C. Ding and H. Peng, "Minimum Redundancy Feature Selection from Microarray Gene Expression Data," 2003. doi: <https://doi.org/10.1142/s0219720005001004>.
15. MathWorks, "Support Vector Machine (SVM)," MathWorks. Accessed: Feb. 18, 2023. [Online]. Available: <https://ww2.mathworks.cn/en/discovery/support-vector-machine.html>
16. MatLab, "Choose Classifier Options." Accessed: Oct. 26, 2023. [Online]. Available: Choose Classifier Options
17. M. Sohaib and J. M. Kim, "Data driven leakage detection and classification of a boiler tube," Applied Sciences (Switzerland), vol. 9, no. 12, Jun. 2019, doi: 10.3390/app9122450.
18. S. Carrino, J. Guerne, J. Dreyer, H. Ghorbel, A. Schorderet, and R. Montavon, "Machining Quality Prediction Using Acoustic Sensors and Machine Learning," MDPI AG, Dec. 2020, p. 31. doi: 10.3390/proceedings2020063031.
19. X. Ying, "An Overview of Overfitting and its Solutions," in Journal of Physics: Conference Series, Institute of Physics Publishing, Mar. 2019. doi: 10.1088/1742-6596/1168/2/022022.
20. D. Normawati and D. P. Ismi, "K-Fold Cross Validation for Selection of Cardiovascular Disease Diagnosis Features by Applying Rule-Based Datamining," Signal and Image Processing Letters, vol. 1, no. 2, pp. 23–35, Jul. 2019, doi: 10.31763/simple.v1i2.3.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

