Received: 16 November 2022 | Revised: 14 January 2023 | Accepted: 15 January 2023
DOI: 10.37917/ijeee.19.1.13 Vol. 19| Issue 1| June 2023

Open Access

Iraqi Journal for Electrical and Electronic Engineering
Original Article

Machine Learning Approach Based on Smart Ball COMSOL Multiphysics Simulation for Pipe Leak Detection

Marwa H. Abed*, Wasan A. Wali, Musaab Alaziz
Department of Computer Engineering, University of Basrah, Basrah, Iraq

Correspondence
*Marwa H. Abed
Department of Computer Engineering,
University of Basrah, Basrah, Iraq
Email: engpg.marwa.abd@uobasrah.edu.iq

Abstract
Due to the changing flow conditions during the pipeline's operation, several locations of erosion, damage, and failure occur. Leak prevention and early leak detection techniques are the best pipeline risk mitigation measures. To reduce detection time, pipeline models that can simulate these breaches are essential. In this study, numerical modeling using COMSOL Multiphysics is suggested for different fluid types, velocities, pressure distributions, and temperature distributions. The system consists of 12 meters of 8-inch pipe. A movable ball with a diameter of 5 inches is placed within. The findings show that dead zones happen more often in oil than in gas. Pipe insulation is facilitated by the gas phase's thermal inefficiency (thermal conductivity). The fluid mixing is improved by 2.5 m/s when the temperature is the lowest. More than water and gas, oil viscosity and dead zones lower maximum pressure. Pressure decreases with maximum velocity and vice versa. The acquired oil data set is utilized to calibrate the Support Vector Machine and Decision Tree techniques using MATLAB R2021a, ensuring the precision of the measurement. The classification result reveals that the Support Vector Machine (SVM) and Decision Tree (DT) models have the best average accuracy, which is 98.8%, and 99.87 %, respectively.

KEYWORDS: Leakage Detection, Fluid Flow, Heat Transfer, Pipeline Monitoring, SVM, DT, Computational Fluid Dynamics.

I. INTRODUCTION

According to the Worldwide World Energy Report in 2020, natural gas and fossil fuels accounted for 30% of the global demand for energy production[1]. Fuel delivery, including fuel supply through fluid pipes, is essential to energy provision [2]. Applied stress, environmental factors, and noise levels are among the many challenges that pipelines encounter, with mechanical systems deteriorating to various extents [3]. Certain factors in the leakage phenomenon are challenging to measure in actuality. Some researchers have recently embraced computational fluid dynamics as a technique to aid in this process. Many industrial operations involve the simultaneous flow of two immiscible liquids in vertical pipes, one of which is the petroleum industry [4]. Because of the importance of the topic, many authors have focused their attention on the study of methods that could be used to identify leaks in pipelines that are used for the production and transit of oil [5]–[7]. At the current time, many different techniques for detecting leaks, including those based on harmful pressure waves, acoustic sensors, satellite surveillance, mass and volume balance, and analytical model-based procedures, have been put into practice. These methods depend on many aspects of the process, such as the temperature, pressure, mass and volumetric flow rates, and so on [8].

According to Dong et al. [9], the most beneficial of these technologies is the negative pressure strategy since it offers great leak sensitivity and availability. Unfortunately, this method has a high likelihood of creating a false alert if the pressure measurement records indicate significant changes or if the leak is tiny (0.5% of nominal flow) [10]. As a result, it needs to investigate using numerical methods into the hydrodynamics of heavy oil-water flow in a vertical pipe with a slight leak, which is much more challenging to detect using conventional methods [11]. Pipelines that are utilized in real-time operations are frequently situated in extreme environments, such as those found in the sea, where they have been subjected to the pressure that is exerted by the water; in the middle of the desert; or even underground, where they are subjected to the force that is exerted by the

force of soil stress [12]. Pipelines utilized in real-time operations are frequently located in these kinds of environments. The most serious issue is the difficulty in discovering leaks caused by physical corrosion and pipes age [13]. The situation deteriorates when used over an extended period without being discovered. Most leaks could be located using the trial-and-error method that involves watching a change in the fluid flow pattern [14]. There may be a sudden change in pressure across the free stream due to the leak, which would cause the flow to be distorted. The performance of systems that include the mobility of tiny spherical particles about the fluids in which they could be immersed involves a wide range of phenomena that are significant to researchers and engineers. The terminal settling velocity of a single spherical particle in an infinite fluid is of interest to various fields, as is the fluid flow drag (pressure drop) that a sphere experiences during fluid flow. Consequently, several empirical and theoretical investigations have been concentrated on them. One may reasonably presume that all significant problems were addressed some decades ago. Significant physical and mathematical ramifications are associated with using an analytical drag formula. The intricacy of the flow prevents an analytical description of the friction factor and viscous forces from being possible across a wide range of velocities where the pressure drop of leakage detection cannot be easier to detect theoretically [15]–[17]. Research on the pressure drop in the sphere area has been undertaken throughout a wide range of flow velocities, fluid viscosity, and density ever since the pioneering investigations of Stokes, Osteen, and others for flow through a ball. The characteristics of the fluid will change in general as a result of the change in the fluid [18], [19]. The recent studies are performed using numerical simulation for leakage detection modeling as seen in Table I in addition to processing the real or numerical data utilizing various machine learning Algorithms.

TABLE I

LEAK DETECTION MODELING IN RECENT STUDIES AND PRESENT WORK SCOPE.

Reference	Scope
Barbosa et al. [20]	Modeling of leakage effect on pressure and velocity profiles of (oil-gas) two-phase flow in a gas pipeline using leak localization (LL) functionality
Sousa et al. [21]	Modeling of leakage behavior on pressure and velocity distribution under two-phase flow (water-oil) conditions in the presence using CFX (ANSYS) software where the Finite Element Method and CFD modeling were applied
Araújo et al., [22]	a numerical investigation of two leakage points on oil pipeline using CFX simulation where the leakage points effect on pressure and velocity profiles were obtained
Adebayo [23]	use data-driven intelligent models of machine learning to find the small leakage depending upon the behavior of flow and heat of the gas pipeline
Sharma [24]	use SVM for leakage detection in the gas pipeline using image processing.
Shen and Cheng [25]	use on-site data for leakage detection utilizing machine learning models in Water Distribution Systems(WDS)
Proposed investigation	- Fluid flow and temperature transfer modeling of various fluids pipelines in presence of leakage detection ball. - Resultant data from numerical simulation using in training of machine learning (SVM and DT) algorithm for realistic leakage detection.

II. MACHINE LEARNING IN PIPELINE LEAK DETECTION

Support Vector Machine (SVM), which is selected since it works well in a wide subspace, can efficiently manage vast volumes of data, and is ideal for categorizing non-linear input, is the algorithm that will be used [26]. A support vector machine is a type of supervised machine learning model that solves issues involving two groups of categorizations by using classification techniques. When an SVM model is provided with sets of labeled training data for each category, the model can classify newly encountered text. When compared to the most recent algorithms, such as neural networks, they offer two primary advantages: increased speed and improved performance with a constrained quantity of data points (in the thousands). Because of this, the approach is well suited for solving issues involving the categorization of text, which often involves having access to datasets containing no more than a few thousand annotated examples at most [13].

Adebayo [23] uses data-driven intelligent models to find small leakage depending upon the behavior of the flow and heat of the gas pipeline. He found that SVM gives a good approximation for leakage detection but does not provide the most accurate measurement and DT is the most sensitive algorithm. Sharma [24] uses SVM for leakage detection in the pipeline using image processing. This approximation gives a superior leakage indication based on practical reality. The image resolution promoted no superior accuracy (15 % tolerance). Shen and Cheng [25] use on-site data for leakage detection utilizing a machine learning algorithm. Adaboost, random forest, and Discussion tree are used in their investigation. DT has the most accurate approximation with a minimum false rate.

This paper proposes techniques to estimate leak location and leakage rate by creating a model to simulate leak detection in the oil transmission pipeline utilizing a movable

ball detecting leaks within the pipeline using acoustic signals. With this information, trained guesses can be made about where the leak is and how fast it is spreading. The velocity, pressure, and temperature profiles will be used to calibrate the internal control system using the Support Vector Machine (SVM) and Decision Tree (DT) algorithms for classification leaks using the ball. The control system factors in the disruptions brought about by the fluid flow around the ball, and it makes a relationship between the levels of sound pressure and the detection of leaks.

Machine learning is a branch of artificial intelligence (AI) and computer science that focuses on using data and algorithms to imitate how humans learn, to constantly improve the simulation's accuracy. Several kinds of machine learning algorithms are often employed. These are as follows: Support Vector Machine (SVM), Decision trees (DT), and K- Nearest-Neighbor (KNN). This paper uses SVM and DT for comparison. The obtained data of velocity, pressure, and temperature distribution parameters, where the leakage sound energy is applied, are employed within the linear regression algorithm using SVM[27]. The detailed steps for developing the current SVM system are illustrated in Fig. 1,

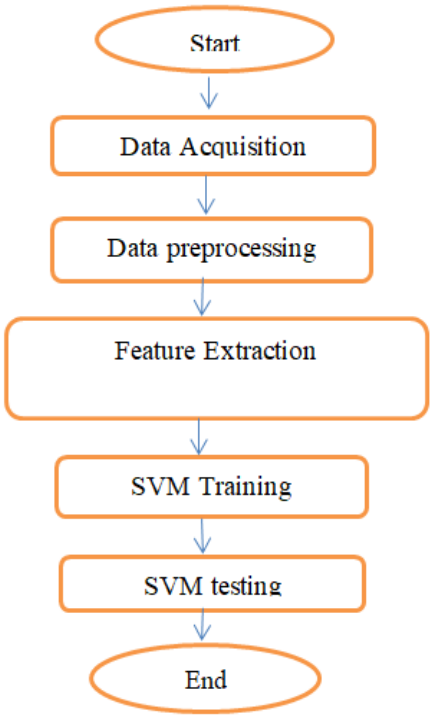


Fig. 1: SVM detailed steps.

MATLAB R2021a is used throughout this study to carry out Support Vector Machine calculations (SVM) where the optimum correction curve is utilized statistically using the following Equation:

$$RMSE = \left(\frac{\sum n(yave-y)^2}{n}\right)^{0.5} \quad (1)$$

Where RMSE is a root mean square deviation of resultant error.

y: is response values

n: is the number of data set.

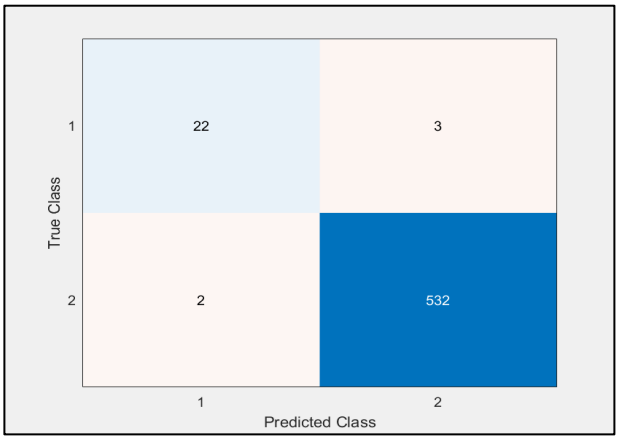
The root-mean-square error (RMSE) is a popularly used measurement of the gaps that exist between the values (both sample and population values) that are forecasted by a model or estimator and the values that are observed. The resultant optimization model is used for measurement modification. When explaining or summarizing the expected results of a classification problem, confusion matrices are a beneficial tool. Linear regression algorithm develops Support Vector machine prevails Confusion matrix. A Confusion matrix's most crucial function is to provide a class-by-class breakdown of the total number of correct and incorrect guesses that have been generated. The performance of the machine learning method can be investigated using the parameters of precisions, accuracy, recall, and F1-score. The confusion matrix determines the following [28], [29]:

- True Positive (TP): Both the value that was seen and the value that was predicted are positive.
- False Negative (FN): When the actual observed value is interpreted as having a negative sign, even when it has a positive one.4
- The condition that is referred to as a “True Negative” (TN) is one in which the observations are consistent with the expectations of the null hypothesis.

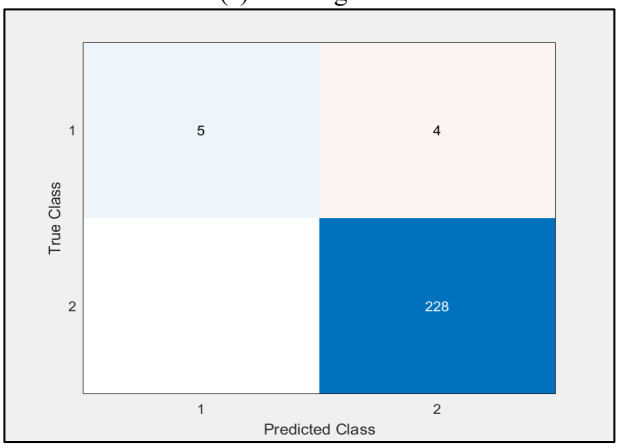
When it comes to classifiers, the Receiver Operating Characteristic (ROC) graph is a useful tool for determining which element is the most essential. The rate of true positives is represented along the ROC curve’s Y axis, while the rate of false positives is shown along the ROC curve’s X axis. The “ideal” location, which may be found at the top left corner of the map, has a failure probability of zero and a success probability of one. It is evident that this is not the case; nevertheless, it does show that a larger Area under the curve (AUC) is desirable in the majority of circumstances [30], [31]. The “steepness” of ROC curves is one factor that may affect the ideal strategy, which is to increase the TP rate while decreasing the FP rate. The analytical result was negative, but the projection is that it will be positive. ROC curves are frequently used for binary classification to evaluate the output of a classification, and this is exactly what is being done here because the classification technique includes whether or not a leakage is identified [31].

The SVM and DT were trained using our dataset generated from simulation, with extra leakage points added for optimum performance during training and testing. The data is divided using cross-validation (K-fold) with K=10, with 70% of the data randomly selected for training and 30% for testing, with the accuracy evaluated at each iteration. The data consist of three parameters (velocity, pressure, and temperature). Each parameter has velocity values (0.1 m/s, 1 m/s, and 2.5 m/s). The confusion matrices of SVM after training and testing the data is shown in Fig. 2, and the confusion matrices of DT for training and testing data is shown in Fig. 3,

As the figures show, the rows of the confusion matrices represent the true classes, whereas the columns represent the predicted classes. The numbers in the blue cells represent correct predictions. These cells have higher values than the other cells in the same row. Therefore, it can be assumed that the model performed well if the diagonal (from top to bottom) was highlighted.

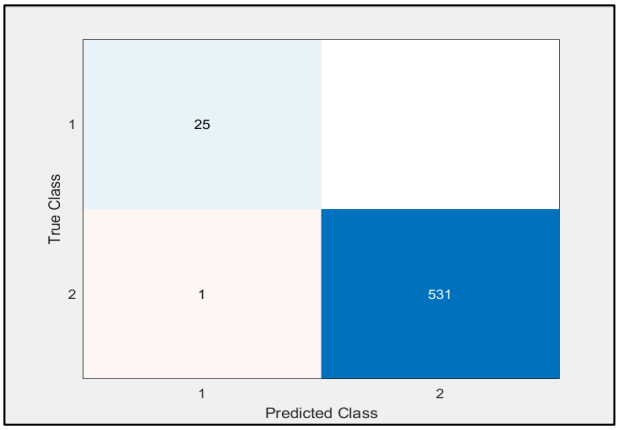


(a) Training data

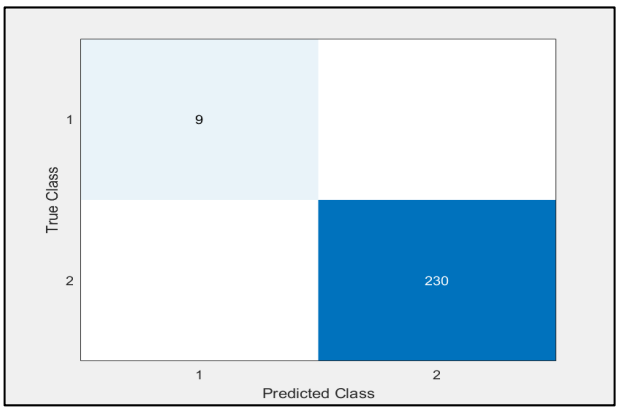


(b) Testing data

Fig. 2: The confusion matrices for SVM



(a) Training data



(b) Testing data

Fig. 3: The confusion matrices for DT

III. Modeling Using Comsol Multiphysics®

The COMSOL Multiphysics® Software simulates the actions of a smart ball with sensors to identify oil/water pipeline breaches. COMSOL Multiphysics is a cross-platform finite element analysis solution. It enables physics-based and linked PDEs (PDEs). It is also monitoring related tasks to improve an organization's performance. Systems engineering, industrial engineering, and validation sciences are related. COMSOL oversees the purpose and status of projects within a program and can use this oversight to support project-level activity to ensure program goals are met by providing decision-making capacity that cannot be achieved at the project level, providing the project manager with a program perspective when needed, or serving as a sounding board for ideas and problem-solving approaches. COMSOL may be well-positioned to deliver this knowledge by aggressively soliciting information from project managers. Electrical, Mechanical, Fluid, Chemical, Multipurpose, and Interfacing modules are available. COMSOL Multiply allows linking and communication between simulations. This horrible scenario is real [3]. COMSOL Multiphysics 5.6 solves fluid, thermal, and mechanical problems. COMSOL Multiphysics model builder is a tree. The COMSOL tree's main components are [32]:

- Geometry builder imports solid work geometry via a graphical interface.
- Material specification: This item modifies the physical characteristics of materials in the COMSOL database or generates new blank properties.
- Physics selection specifies which physics will be utilized to apply conservation equations and how they will be connected and regulated. Physics needs the boundary condition to solve.
- COMSOL's mesh generation is based on physics and geometry. Normal, coarse, coarser, extremely coarse, fine, finer, and very fine.
- Stable (steady state) or time-dependent model-solving study (unsteady state).

COMSOL Multiphysics was used to develop the pipeline form of this mobile inspection ball. The COMSOL Computational Fluid Dynamics (CFD) module was used to

describe the velocity and pressure propagation around the ball in the pipeline and connect it to heat transfer to compute the temperature around the pipes and its transfer to the liquid in two cases:

I) The pipeline's continuous flow of fluid with no leaks.
II) The pipeline leaks a ball detector.

The CFD module calculates ball temperature by analyzing velocity and pressure characteristics. Simulate leak-caused noise using Acoustics and Vibrations. COMSOL Multiphysics has a tree-based model builder.

COMSOL Multiphysics was used to design this self-driving ball's pipeline shape. The COMSOL CFD module was used to model the velocity and pressure propagation around the ball inside the pipeline. Then it was linked to heat transfer to compute the temperature around the tubes and its transmission to the liquid.

The ball's velocity, pressure, and temperature are all determined via the CFD module. The next step is to replicate the noise that is created by a leak using the acoustics and vibrations module. COMSOL Multiphysics was used to build the pipeline form of this autonomous ball, which represents a novel design in mobile inspection equipment. The ball itself is a revolutionary design. The COMSOL Computational Fluid Dynamics (CFD) module was then used to model the velocity and pressure propagation around the ball positioned inside the pipeline.

The CFD module is used to identify the velocity and pressure profiles located in the area around the moving ball, in addition to the heat transfer used to compute the temperature. The next step is to use the acoustics and vibrations module to simulate the spread of noise created when a leak occurs. The model under investigation comprises the flow of fluid around a stationary spherical ball placed within a pipeline, as well as the propagation of sound created by an induced leak. To get an appropriate comprehension of the collected data, Multiphysics was included in the construction of the model. Fig. 4, depicts an algorithm.

The use of single-phase laminar flow, where μ is dynamic viscosity, ρ is density, U is used velocity, and F is applied momentum force due to pressure sound at a specified point. The continuity equation (1), consists of the conservation of control volume mass with inflow and outflow transported mass. Navier Stoke equation (2), refers to the equality of inertia term to the sum of viscous force, gravity force, and volume force due to the acoustic action.

- Fluid flow equations

$$\boldsymbol{\rho}\left(\boldsymbol{U_r}\frac{\partial}{\partial \boldsymbol{r}}(\boldsymbol{u_r})+\frac{\boldsymbol{u_\theta}}{\boldsymbol{r}}\frac{\partial}{\partial \boldsymbol{\theta}}(\boldsymbol{u_r})+\boldsymbol{u_z}\frac{\partial U_r}{\partial \boldsymbol{z}}\right)=-\frac{\boldsymbol{dp}}{\boldsymbol{dr}}+\boldsymbol{F_r}+\boldsymbol{\mu}\left(\frac{1}{\boldsymbol{r}}\frac{\partial \boldsymbol{u_r}}{\partial \boldsymbol{r}}+\frac{1}{\boldsymbol{r^2}}\frac{\partial^2 \boldsymbol{u_r}}{\partial \boldsymbol{\theta^2}}+\frac{\partial^2 \boldsymbol{u_r}}{\partial \boldsymbol{z^2}}\right) \quad (2)$$

$$\boldsymbol{\rho}\left(\mathbf{U_r}\frac{\partial}{\partial \mathbf{r}}(\mathbf{u_\theta})+\frac{\mathbf{u_\theta}}{\mathbf{r}}\frac{\partial}{\partial \boldsymbol{\theta}}(\mathbf{u_\theta})+\mathbf{u_z}\frac{\partial \mathbf{u_\theta}}{\partial \mathbf{z}}\right)=-\frac{\mathbf{dp}}{\mathbf{d\theta}}+\mathbf{F_\theta}+\boldsymbol{\mu}\left(\frac{1}{\mathbf{r}}\frac{\partial \mathbf{u_\theta}}{\partial \mathbf{r}}+\frac{1}{\mathbf{r^2}}\frac{\partial^2 \mathbf{U u_\theta}}{\partial \boldsymbol{\theta^2}}+\frac{\partial^2 \mathbf{u_\theta}}{\partial \mathbf{z^2}}\right) \quad (3)$$

$$\boldsymbol{\rho}\left(\mathbf{U_r}\frac{\partial}{\partial \mathbf{r}}(\mathbf{u_z})+\frac{\mathbf{u_\theta}}{\mathbf{r}}\frac{\partial}{\partial \boldsymbol{\theta}}(\mathbf{u_z})+\mathbf{u_z}\frac{\partial \mathbf{u_z}}{\partial \mathbf{z}}\right)=-\frac{\mathbf{dp}}{\mathbf{dz}}+\mathbf{F_z}+\boldsymbol{\mu}\left(\frac{1}{\mathbf{r}}\frac{\partial \mathbf{u_z}}{\partial \mathbf{r}}+\frac{1}{\mathbf{r^2}}\frac{\partial^2 \mathbf{U u_z}}{\partial \boldsymbol{\theta^2}}+\frac{\partial^2 \mathbf{u_z}}{\partial \mathbf{z^2}}\right) \quad (4)$$

- Heat transfer equation

$$\boldsymbol{\rho}\left(\mathbf{U_r}\frac{\partial}{\partial \mathbf{r}}(\mathbf{T})+\frac{\mathbf{u_\theta}}{\mathbf{r}}\frac{\partial}{\partial \boldsymbol{\theta}}(\mathbf{T})+\mathbf{u_z}\frac{\partial \mathbf{T}}{\partial \mathbf{z}}\right)=\mathbf{Q}+\mathbf{K}\left(\frac{1}{\mathbf{r}}\frac{\partial \mathbf{T}}{\partial \mathbf{r}}+\frac{1}{\mathbf{r^2}}\frac{\partial^2 \mathbf{T}}{\partial \boldsymbol{\theta^2}}+\frac{\partial^2 \mathbf{T}}{\partial \mathbf{z^2}}\right) \quad (5)$$

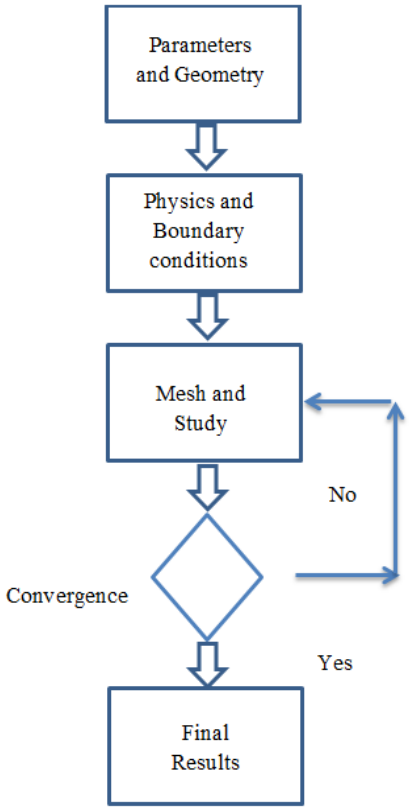


Fig. 4: The algorithm steps of solving the current model.

Initial conditions for laminar flow are v=0 m/s and p=0 Pa, while boundary conditions include an inlet, exit, wall, and pressure point limitations owing to leaks. Initial heat transfer conditions are 293.15 K, and boundary conditions include fluid, thermal insulation, solid domain heat transfer, inflow, outflow, heat flux, and point heat source. Initial acoustic dispersion is at w=0 j/m3. The boundary condition is a leaky powerpoint. Reference temperature (293 K) establishes fluid flow physical attributes (viscosity and pressure) and density distribution within the pipe. First, the fluid flow (momentum transfer) is solved using the Navier Stoke equation [10]. The thermal physical parameters (thermal conductivity and heat capacity) are generated by reference temperature to solve the heat transfer equation with the momentum transport equation; then the temperature distribution is gentle. The trial-and-error approach loops are used based on the first phase where a new temperature distribution replaces a reference temperature, etc. New temperature implies new physical characteristics and velocity distribution; the cycle is continued until the error residence is 0.1% or less. The acoustic pressure point creates non-uniform distribution in the free stream as an extra sound power source within the momentum force.

To simulate the fluid flow around the ball, a section of a typical cylindrical pipeline model was developed having a

10-inch pipe length of 12 m, as shown in Fig. 5, Oil, water, and gas were used as the fluid flow, which was given an initial inlet of three velocities 0.1, 1, 2.5 m/s, respectively. The ball has a diameter of 5 inches. Fig. 6, shows the mesh distribution of the proposed system. The present problem is solved by utilizing the 96245 mesh elements; all of them are tetrahedral. The simulation generates 96245 algebraic square matrices from utilized atrial differential equations, all of which are solved using numerical methods, such as Jacobean.

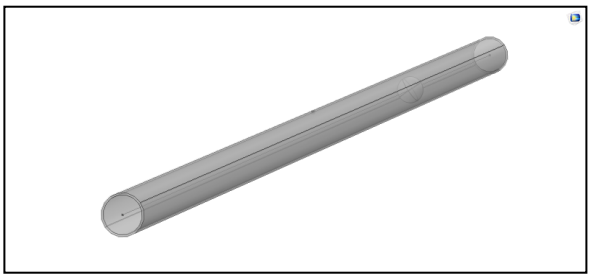


Fig. 5: The geometrical view of a spherical ball inside the pipe.

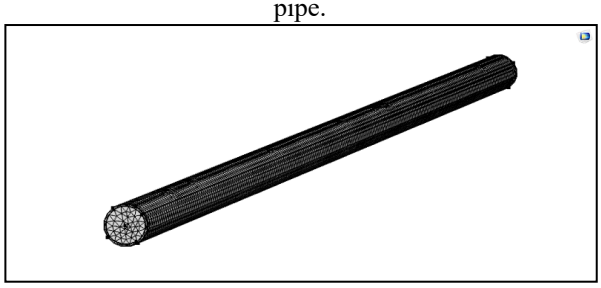
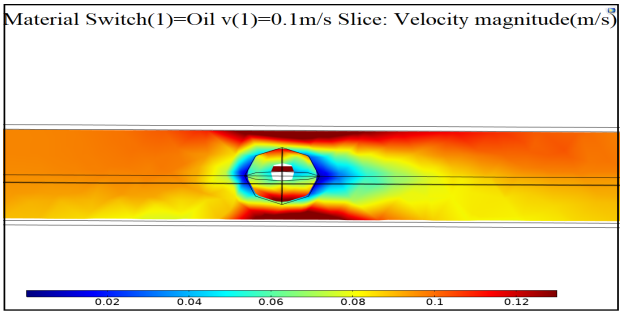


Fig. 6: Mesh distribution of the present system.

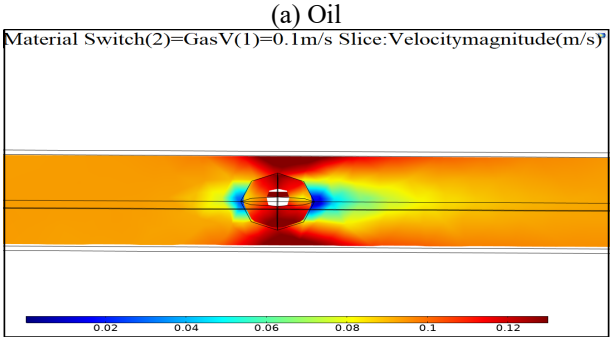
IV. Results And Discussion

A. Numerical Results

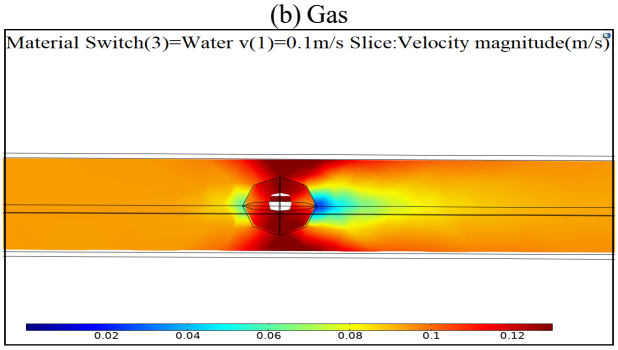
This simulation used the Computational Fluid Dynamics (CFD) module. Fig. 7, shows the velocity contours for various fluids. The dead flow zones (blue color regions) are formed in solid wall regions (pipe wall and ball wall). The oil has the maximum dead zones while the gas has a minimum. The dead zones are a boundary layer formed by the mean of viscous forces Fig. 8, shows the velocity distribution between water, gas, and oil for various velocities. The fluid type has no significance on velocity distribution, and the ball region prevails over minimum velocity values. The oil has lower velocity values than the water case; the viscosity of the oil is higher than water, indicating the flow resistance behavior. Natural gas has higher velocity values than water, yet its density is lower, resulting in fewer momentum forces in the pipeline system



(a) Oil

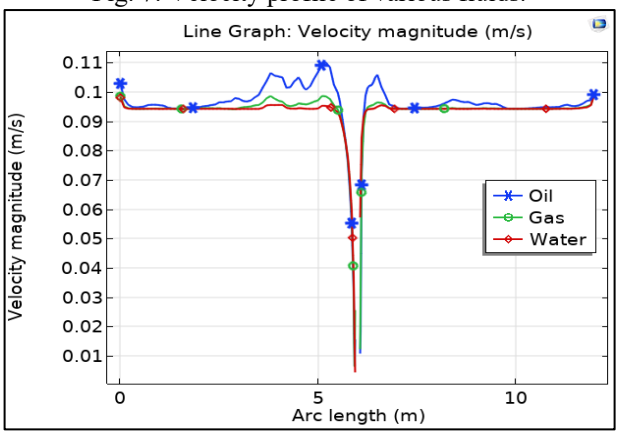


(b) Gas



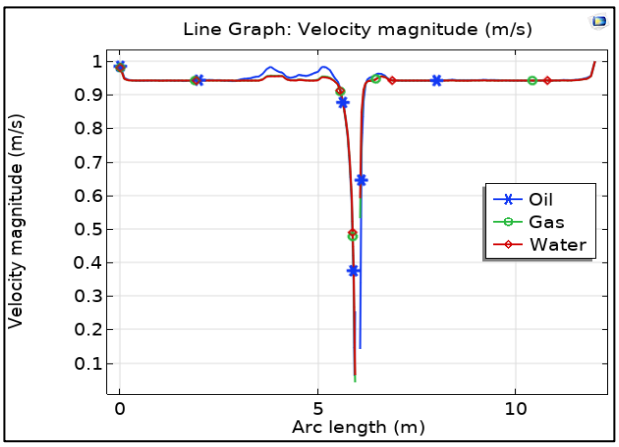
(c) Water

Fig. 7: Velocity profile of various fluids.

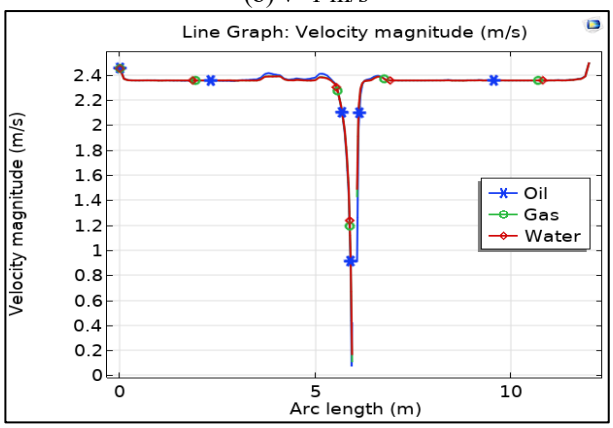


(a) V=0.1 m/s

Fig. 8: Comparison of velocity distribution between water, gas, and oil for various velocities.



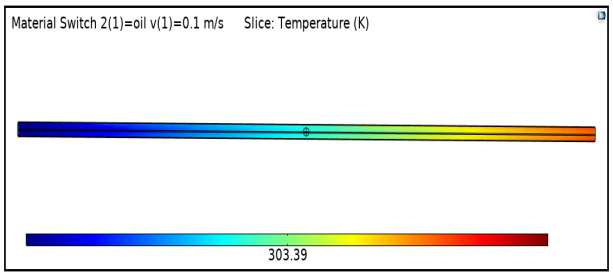
(b) v=1 m/s



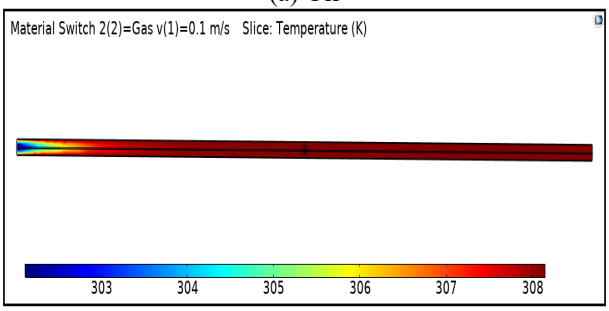
(c) v=2.5 m/s

Fig. 8: Continued.

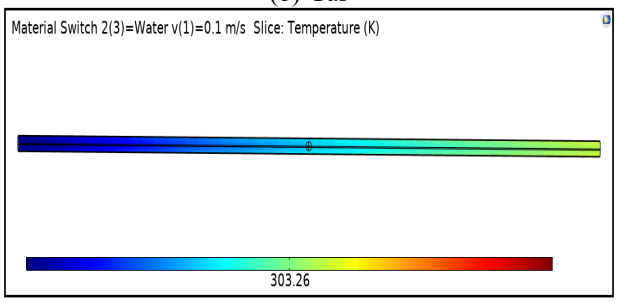
Fig. 9, shows the temperature contours for various fluids. The gas fluid has a maximum temperature gradient, while the minimum water gradient is observed in the water case. The thermal conductivity of gas is less than oil, and water has maximum thermal conductivity. The gas acts as thermal insulation from the pipe wall and bulk fluid. The leakage can be detected by the mean of temperature distribution easily. Fig. 10, shows the comparison of temperature distribution between water, gas, and oil for various velocities. The temperature distribution of gas is the minimum value for whole inlet velocity values. The physical properties, especially the thermal characteristics of the gas, are less than oil and water, such as thermal conductivity. The heat energy due to pressure sound is restricted by the thermal properties of the used fluid. The higher temperature difference between the inflow and outflow of the pipeline in the natural gas case indicates that the physical properties (thermal properties) of natural gas are less than the liquids. The temperature difference in the gas case decreases as inlet velocity increases. Fig. 11, shows the comparison of the temperature distribution in gas for various velocities. The maximum velocity (2.5 m/s) provides a minimum temperature gradient, and the temperature gradient decreases as velocity increases. The increasing velocity enhances the cold and hot fluid particles, maximizing heat transfer tendency.



(a) Oil



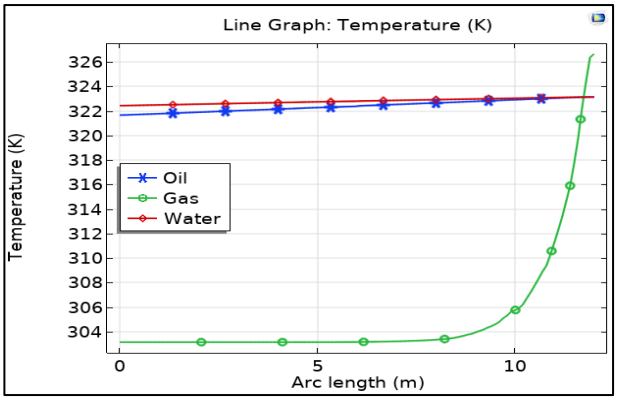
(b) Gas



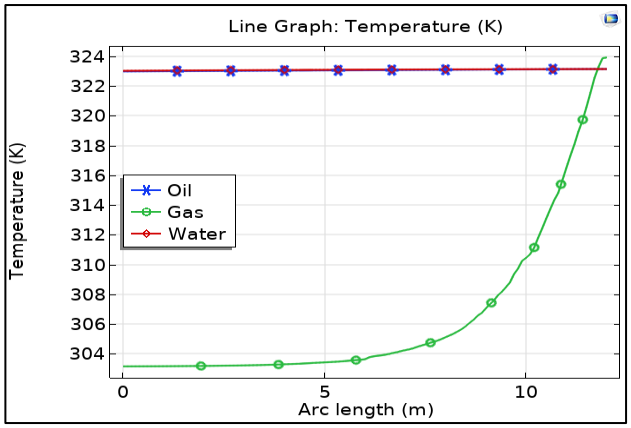
(c) Water

Fig. 9: Velocity profile of various fluids.

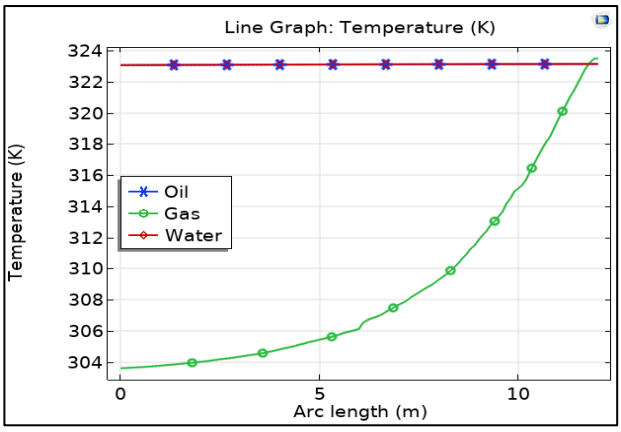
Fig. 12, shows the comparison in pressure distribution between water, gas, and oil for various velocities. The pressure distribution of oil is the maximum value for whole inlet velocity values. The viscous forces, due to dead zones, can be exhibited as pressure drop. The maximum pressure drop indicates boundary layer development where the laminar flow is performed. Fig. 13, shows the comparison in pressure distribution in gas for various velocities. The increase in velocity increases pressure values. The velocity component converts into pressure forces using energy conservation. Fig. 14, illustrates a Pressure comparison between leaks, non-leaks, and non-leak-non-ball cases where oil flows in 2.5 m/s. The leakage case promotes negative pressure values near the pressure region while the whole pressure values are positive in the absence of leakage. Leaks can easily cause sudden changes in pressure distribution.



(a) V=0.1 m/s



(b) V=1 m/s



(c) V=2.5 m/s

Fig. 10: The comparison in Temperature distribution between water, gas, and oil for various velocities.

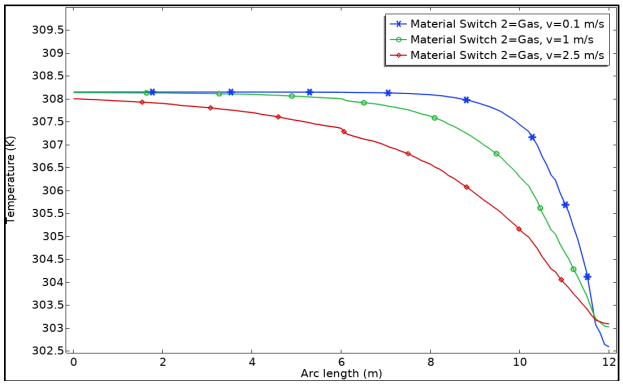
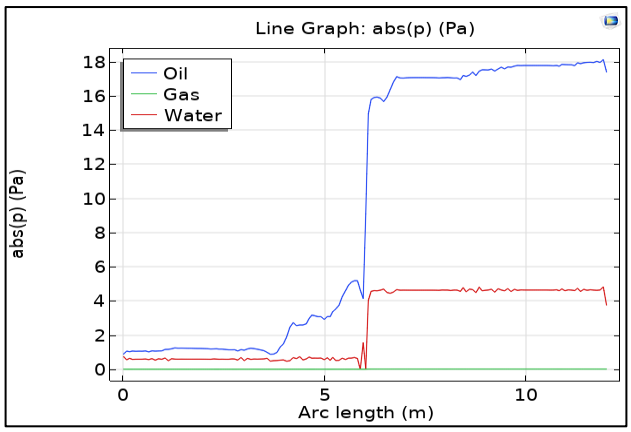
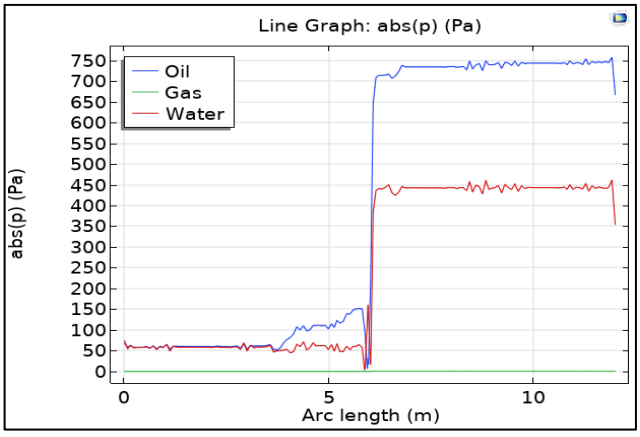


Fig. 11: Temperature distribution of various velocities; gas is used as the fluid.

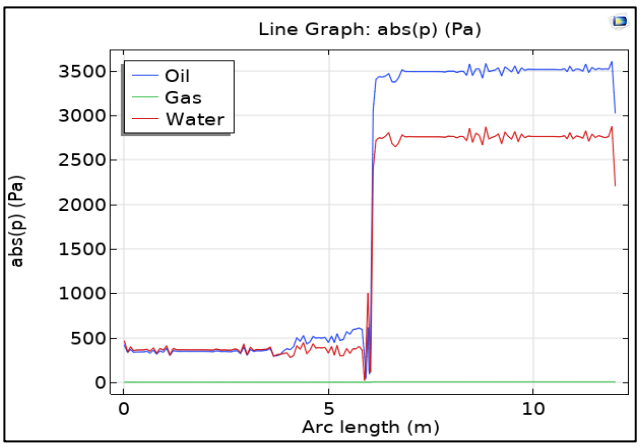


(a) V=0.1 m/s



(b) V= 1 m/s

Fig. 12: The comparison in pressure distribution between water, gas, and oil for various velocities.



(c) V =2.5 m/s

Fig. 12: Continued.

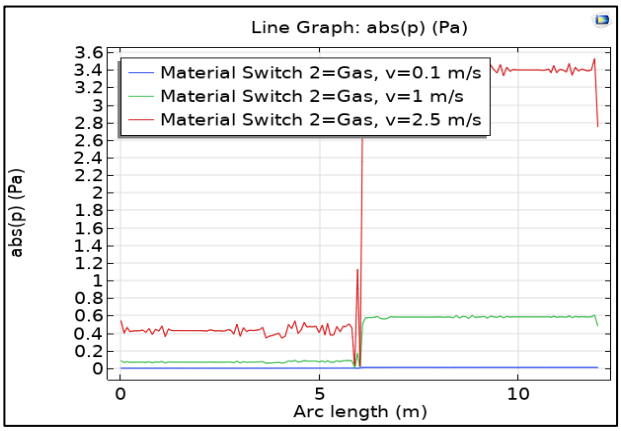


Fig. 13: Pressure distribution of various velocities, gas is used fluid.

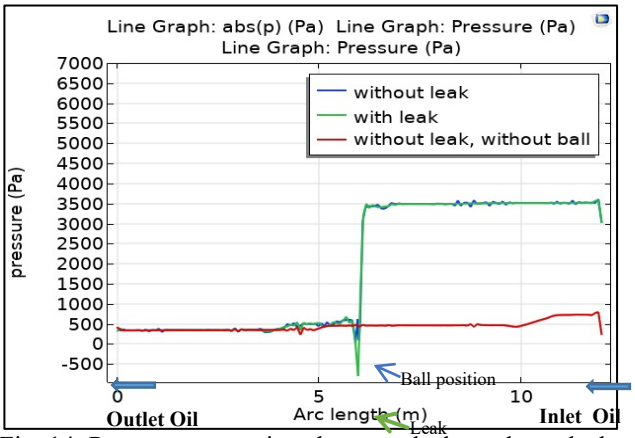


Fig. 14: Pressure comparison between leaks and non-leaks cases where oil flows in 2.5 m/s.

B. *Machine learning models results*

The training phase of SVM and DT is divided into 70% for training and 30% for testing by cross-validation, and the average accuracy is 98.8%, and 99.87%, respectively. Table II shows the SVM and DT model's average precision, recall, and F1-score of the present work. The DT has perfect precision, Recall, and F1-score as compared with SVM. The membership function of DT is more than SVM, the DT algorithm is more complicated. The DT is considered a confirmation algorithm more than SVM because it does not deal with the dependent and independent data as linear or non-linear regression. While SVM should be specified for linear or nonlinear expressions which must be solved by Gaussian approximation [33], the accuracy is determined by the following equation:

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (6)$$

The following equations also determine the precision, recall, and F1-score for the SVM model:

$$Precision = \frac{TP}{TP+FP} \quad (7)$$

$$Recall = \frac{TP}{TP+FN} \quad (8)$$

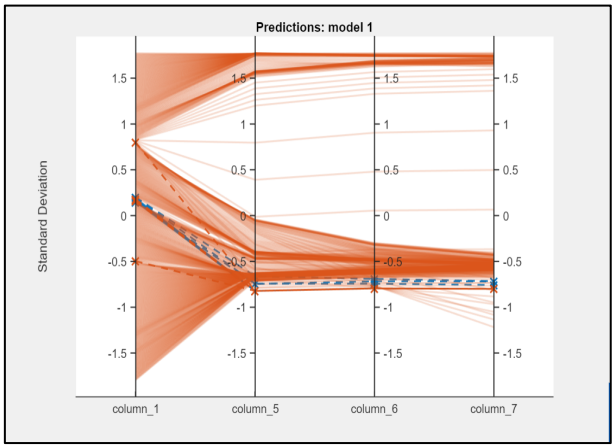
$$F1-score = 2 \times \frac{Precision \times Recall}{Precision+Recall} \quad (9)$$

Where:
TP is True Positive
TN is True Negative
FP is False Positive, and
FN is False Negative

TABLE II
Precision, recall, and f1-score for the (SVM) and (DT) models.

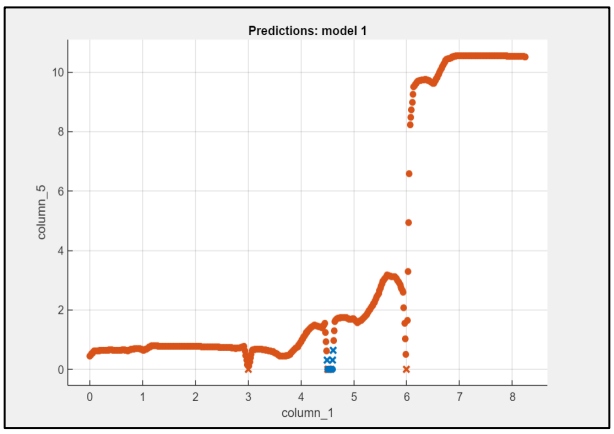
Model	Precision %	Recall %	F1-score %
SVM	91.67	88	89.8
DT	100	97	100

Fig. 15, shows that the optimization plots are developed based on trained values; the parallel coordinate of column interactions (position, pressure at 0.1 m/s, pressure at 1 m/s, and pressure at 2.5 m/s). The scatter plot indicates the pressure distribution of the leaked and non-leaked points are presented based on optimization plots and confusion matrix.



(a) Parallel coordinates of optimization.

Fig. 15: The optimization plot and Scattering examination using the SVM model.



(b) Scattering plot examination.

Fig. 15: Continued.

V. Conclusion

The results of a numerical examination of the velocity, pressure, and temperature distribution have been satisfactorily modeled over various velocities and kinds of fluid. In the case of oil, the production of the most dead zones is seen, whereas, in the case of gas, the values are at their lowest. Due to the thermally poor qualities of the gas phase, insulation activity is encouraged to take place within the pipe (thermal conductivity). The 2.5 meters per second velocity enables more fluid mixing in the region where the temperature distribution is most uniform. Compared to water and gas, oil has a higher maximum pressure drop because of its viscous forces and dead zones. The highest pressure drop and highest velocity are both provided by the maximum velocity. In this paper, An SVM was proposed for detecting leakage at four points in 3 m, 4.5 m, 6 m, and 9 m pipelines using a pressure sensor. The dataset of this study was obtained from the simulation with three parameters of velocity, pressure, and temperature. Each parameter consists of three values (0.1, 1, 2.5 m/s). The SVM and DT models get an average accuracy (validation) of 98.8%, and 99.87%, respectively for classification. Comparing the results obtained, we conclude that we obtain higher accuracy when training and testing the DT. So, for this data type, it is best to use the DT algorithm because it is faster and more accurate than another algorithm. For future work, different classification algorithms could be applied and the resulting classification accuracy could be compared to achieve higher efficiency.

Conflict of Interest

The authors have no conflict of relevant interest to this article.

References

[1] R. Kroes, L. Amundsen, and R. Hoffmann, "Crude oil-water flow in horizontal pipes," 2013. http://essay.utwente.nl/69533/

[2] W. A. Wali, "Carbon Dioxide Conversion Control Based on Microwave Plasma Technology," *Int. Conf. Electr. Commun. Comput. Eng.*, pp. 1–4, 2020. doi: 10.1109/ICECCE49384.2020.9179393.

[3] W. R. Chalgham, A. C. Seibi, and M. Mokhtari, "Simulation of Sound Wave Propagation Inside a Spherical Ball Submerged in a Pipeline," *Proc. 2016 COMSOL Conf. Bost.*, p. 2016, 2016.

[4] M. M. Hathal, "Oil fouling in double - pipe heat exchanger under liquid - liquid dispersion and the influence of copper oxide nanofluid," Heat Transfer, vol. 48, Issue 5, pp. 1963-1981, July 2019. doi: 10.1002/htj.21473.

[5] J. V. N. de Sousa, C. H. Sodré, A. G. B. de Lima, and S. R. de F. Neto, "Numerical Analysis of Heavy Oil-Water Flow and Leak Detection in Vertical Pipeline," *Adv. Chem. Eng. Sci.*, vol. 03, no. 01, pp. 9–15, 2013. doi: 10.4236/aces.2013.31002.

[6] C. W. Liu, Y. X. Li, Y. K. Yan, J. T. Fu, and Y. Q. Zhang, "A new leak location method based on leakage acoustic waves for oil and gas pipelines," *J. Loss Prev. Process Ind.*, vol. 35, pp. 236–246, May 2015. doi: 10.1016/j.jlp.2015.05.006.

[7] S. I. Kam, "Mechanistic modeling of pipeline leak detection at fixed inlet rate," *J. Pet. Sci. Eng.*, vol. 70, no. 3–4, pp. 145–156, Feb. 2010. doi: 10.1016/J.PETROL.2009.09.008.

[8] S. Valizadeh, B. Moshiri, and K. Salahshoor, "Multiphase pipeline leak detection based on fuzzy classification," *AIP Conf. Proc.*, vol. 1159, no. 2009, pp. 72–80, 2009. doi: 10.1063/1.3223958.

[9] L. Dong, S. Chai, and B. Zhang, "Leak detection and localization of gas pipeline system based on wavelet analysis," *Proc. 2nd Int. Conf. Intell. Control Inf. Process. ICICIP 2011*, no. PART 1, pp. 478–483, 2011. doi: 10.1109/ICICIP.2011.6008290.

[10] R. Hu, H. Ye, G. Wang, and C. Lu, "Leak detection in pipelines based on PCA," *2004 8th Int. Conf. Control. Autom. Robot. Vis.*, vol. 3, pp. 1985–1989, 2004. doi: 10.1109/ICARCV.2004.1469466.

[11] C. Verde, "Multi-leak detection and isolation in fluid pipelines," *Control Eng. Pract.*, vol. 9, no. 6, pp. 673–682, Jun. 2001. doi: 10.1016/S0967-0661(01)00026-0.

[12] M. Francis, A. Ross, and A. James, "Experimental Observation And Analysis Of Inverse Transient For Pipeline Leak Detection," *ASCE*, vol. 133, no. March, pp. 519–530, 2014.

[13] Y. Liu, X. Ma, Y. Li, Y. Tie, Y. Zhang, and J. Gao, "Water pipeline leakage detection based on machine learning and wireless sensor networks," *Sensors (Switzerland)*, vol. 19, no. 23, pp. 1–21, 2019. doi: 10.3390/s19235086.

[14] U. Baroudi, A. A. Al-Roubaiey, and A. Devendiran, "Pipeline Leak Detection Systems and Data Fusion: A Survey," *IEEE Access*, vol. 7, pp. 97426–97439, 2019. doi: 10.1109/ACCESS.2019.2928487.

[15] Y. Song and S. Li, "Gas leak detection in galvanised steel pipe with internal flow noise using convolutional neural network," *Process Saf. Environ. Prot.*, vol. 146, pp. 736–744, 2021. doi: 10.1016/j.psep.2020.11.053.

[16] A. Terfous, A. Hazzab, and A. Ghenaim, "Predicting the drag coefficient and settling velocity of spherical

particles," *Powder Technol.*, vol. 239, pp. 12–20, May 2013. doi: 10.1016/J.POWTEC.2013.01.052.
[17] M. D. Mikhailov and A. P. S. Freire, "The drag coefficient of a sphere: An approximation using Shanks transform," *Powder Technol.*, vol. 237, pp. 432–435, Mar. 2013. doi: 10.1016/J.POWTEC.2012.12.033.
[18] N. S. Cheng, "Comparison of formulas for drag coefficient and settling velocity of spherical particles," *Powder Technol.*, vol. 189, no. 3, pp. 395–398, Feb. 2009. doi: 10.1016/J.POWTEC.2008.07.006.
[19] R. Barati, S. A. A. S. Neyshabouri, and G. Ahmadi, "Development of empirical models with high accuracy for estimation of drag coefficient of flow around a smooth sphere: An evolutionary approach," *Powder Technol.*, vol. 257, pp. 11–19, May 2014. doi: 10.1016/J.POWTEC.2014.02.045.
[20] A. B. Figueiredo, R. M. Baptista, F. B. de F. Rachid, and G. C. R. Bodstein, "A straightforward strategy for leak localization in two-phase gas pipelines," *J. Nat. Gas Sci. Eng.*, vol. 94, Oct. 2021. doi: 10.1016/j.jngse.2021.104061.
[21] J. V. N. de Sousa, C. H. Sodré, A. G. B. de Lima, and S. R. de F. Neto, "Numerical Analysis of Heavy Oil-Water Flow and Leak Detection in Vertical Pipeline," *Adv. Chem. Eng. Sci.*, vol. 03, no. 01, pp. 9–15, 2013. doi: 10.4236/aces.2013.31002.
[22] M. De Vasconcellos Araújo, S. R. De Farias Neto, A. G. B. De Lima, and F. Daylane Tavares De Luna, "Hydrodynamic study of oil leakage in pipeline via CFD," *Adv. Mech. Eng.*, vol. 2014, 2014. doi: 10.1155/2014/170178.
[23] O. Akinsete and A. Oshingbesan, "Leak detection in natural gas pipelines using intelligent models," *Soc. Pet. Eng. - SPE Niger. Annu. Int. Conf. Exhib. 2019. NAIC 2019*, no. April, 2019, doi: 10.2118/198738-MS.
[24] R. R. Sharma, "Gas Leakage Detection in Pipeline by SVM classifier with Automatic Eddy Current based Defect Recognition Method," *J. Ubiquitous Comput. Commun. Technol.*, vol. 3, no. 3, pp. 196–212, 2021. doi: 10.36548/jucct.2021.3.004.
[25] Y. Shen and W. Cheng, "A Tree-Based Machine Learning Method for Pipeline Leakage Detection," *Water (Switzerland)*, vol. 14, no. 18, 2022. doi: 10.3390/w14182833.
[26] H. Salim, M. Alaziz, and T. Abdalla, "Human Activity Recognition Using The Human Skeleton Provided by Kinect," *Iraqi J. Electr. Electron. Eng.*, vol. 17, no. 2, pp. 183–189, 2021. doi: 10.37917/ijeee.17.2.20.
[27] N. Mashhadi, I. Shahrour, N. Attoue, J. El Khattabi, and A. Aljer, "Use of machine learning for leak detection and localization in water distribution systems," *Smart Cities*, vol. 4, no. 4, pp. 1293–1315, 2021. doi: 10.3390/smartcities4040069.
[28] J. Kemba, K. Gideon, and C. N. Nyirenda, "Leakage detection in Tsumeb east water distribution network using EPANET and support vector regression," *2017 IST-Africa Week Conf. IST-Africa 2017*, Nov. 2017. doi: 10.23919/ISTAFRICA.2017.8102401.
[29] A. Krishnakumari, A. Elayaperumal, M. Saravanan, and C. Arvindan, "Fault diagnostics of spur gear using decision tree and fuzzy classifier," *Int. J. Adv. Manuf. Technol. 2016 899*, vol. 89, no. 9, pp. 3487–3494, Aug. 2016. doi: 10.1007/S00170-016-9307-8.
[30] V. N. Vapnik, "The Nature of Statistical Learning Theory," *Nat. Stat. Learn. Theory*, 2000. doi: 10.1007/978-1-4757-3264-1.
[31] V. Franc and V. Hlaváč, "Vector machine multi-class support," *Proc. - Int. Conf. Pattern Recognit.*, vol. 2, pp. 236–239, 2002. doi: 10.1109/ICPR.2002.1048282.
[32] Karim Egab *et al.*, " Study the Effect of Heat Transfer Coefficient and Thermal Conductivity on Cracked Pipes Carrying Pressurized Fluid," *Int. J. Eng. Technol.* , vol. 8, no. 2019, pp. 275–282, 2019.
[33] M. Ghasemi, M. Zarei, A. Foroutannia, and S. Jafari, "Study of functional connectivity of central motor system in Parkinson's disease using copula theory," *Biomed. Signal Process. Control*, vol. 65, no. November 2020, pp. 102320, 2021. doi: 10.1016/j.bspc.2020.102320.