Journal of Information Systems and Informatics
Vol. 6, No. 1, March 2024 e-ISSN: **2656-4882** p-ISSN: **2656-5935**
DOI: 10.51519/journalisi.v6i1.652 Published By DRPM-UBD

Oil and Gas Pipeline Leakage Detection using IoT and Deep Learning Algorithm

Anietie P. Ekong[1], Gabriel Gregory James[2], Ifeoma Ohaeri[3]

[1]Department of Computer Science, Akwa Ibom State University, Ikot Akpaden, Nigeria
[2,3]Department of Computing, Faculty of Computing and Applied Sciences, Topfaith University, Mkpatack, Nigeria
Email: anietieekong@aksu.edu.ng[1], g.james@topfaith.edu.ng[2], i.ohaeri@topfaith.edu.ng[3]

Abstract

Pipeline leaks are a frequent occurrence in oil and gas infrastructure worldwide. Though leak detection systems are expected to be installed on all pipelines in the near future, relying on human efforts to physically monitor these pipelines is and will continue to be challenging. Though today's leak detection techniques are not able to completely stop leaks from occurring or to detect most leaks, they are essential in lessening their effects. Despite recent developments toward solving this problem, the solution still falls short of expectations. This research presents an approach to pipeline leak detection by leveraging on the exceptional abilities of Convolutional Neural Network (CNN) and Internet of Things (IoT). A comprehensive dataset on oil and pipeline leakage is collected, and the CNN model is developed and trained with the collected dataset. Thereafter, the trained model is integrated into the monitoring system to provide notifications of leaks. The model is adaptable and scalable and its performance, as evaluated, shows an improvement over existing systems with an accuracy of 97% hence well suited for deployment in various pipeline networks for the overall improvement of safety environment in the oil and gas sector.

Keywords: Artificial Intelligence, Convolutional Neural Networks, Computer Vision, Pipeline Leak Detection, Machine Learning

1. INTRODUCTION

The oil and gas industry forms the backbone of modern economies, providing the essential energy resources required for industrial, commercial, and domestic activities. The detection of anomalies is necessary for the smooth operations of oil pipelines and the conventional machine learning binary classification may no longer be sufficient for identifying some of these anomalies like leakages, especially in complex operational conditions [1]. Nevertheless, pipeline leaks are persistent concern when it comes to the movement of fluid or gas across vast networks of pipes. These leaks can lead to catastrophic environmental pollution, economic losses, and potential safety hazards. Despite the existence of conventional leak detection methods, their limitations in terms of real-time responsiveness, accuracy, and adaptability necessitate the exploration of some

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

other solutions such as the machine learning approach, which is research is proffering.

Leakage detection systems should be able to detect spills of different sizes, among other important parameters [2]. To mitigate the losses due to gas pipeline leaks, a digital model based on the signal generated by pressure signals by a pipeline leak could be established [3]. Pipeline leakage incidents can create serious issues to the extent of impeding or attempting to prevent new pipelines as happened in North America [4]. The problem at hand is the inadequacy and efficacy of current pipeline leak detection methods to swiftly identify and localize leaks, particularly in real-time scenarios. Traditional approaches often rely on intermittent inspections or simplistic sensor-based alarms, leading to delayed detection, imprecise location pinpointing, and a higher probability of false alarms. The complexity and dynamic nature of pipeline systems make this issue worse by necessitating a more advanced and flexible method to reliably distinguish between real leaks and innocuous variations in operating parameters.

As pipelines traverse diverse different terrains and climatic zones with their attendant effects on pressures, temperatures, and flow rates, the ability to swiftly and accurately detect leaks becomes an imperative challenge. Addressing the limitations of current leak detection methodologies some of which are manual, partially automated or inconsistent, is paramount to ensuring the safety, environmental sustainability, and operational efficiency of these pipeline networks. This calls for the exploration and implementation of advanced technologies, such as machine learning models, to revolutionize the way leaks are detected and responded to in real-time scenarios. This work aims at using machine learning approach to carry out real-time identification of the oil and gas pipeline leakages.

Machine learning algorithms could be leveraged to ensure that the classification can be done both correctly and timeously and have an excellent ability to handle intricate datasets and perform classification with a high degree of accuracy [5], [6]. Though Conventional Artificial Neural Network like Forward Networks can be used, their predictive ability is a limiting factor. This study holds significance on various fronts, offering profound implications for the oil and gas sector and wider technological advancement. It promises to curtail environmental impact by swiftly pinpointing oil and gas leaks through a Deep learning algorithm. Additionally, operational efficiency could be enhanced by minimizing downtime, lowering maintenance costs, and mitigating resource wastage. Safety standards stand to improve as timely leak identification reduces potential hazards for workers and communities. Adopting the real-time leak detection system could enhance a company's competitive edge by demonstrating environmental responsibility, safety, and operational excellence. Moreover, the study will

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

contribute to the evolving discourse on real-time anomaly detection in complex systems. Regulatory compliance could be streamlined, aligning with stringent environmental and safety regulations. By early detection of leaks, the study supports resource sustainability and the industry's long-term viability. Beyond oil and gas, the methodology has potential applications in other domains with critical infrastructure. Overall, the study's significance lies in reshaping pipeline management, fostering environmental preservation, operational efficiency, safety, technological innovation, and industry competitiveness.

Conventional means to detect leakage in pipes require some degree of expertise by leakage exploration experts and machine learning algorithms have found several applications in pipeline leakage detection, helping to improve the accuracy and efficiency of leak detection systems. Various approaches to machine learning, like the k-nearest-neighbor, random forest models, decision trees, and neural networks have been used to improve detection accuracies [7].

Using the Acoustic Emission (AE) sensor channel information, a platform for leakage identification using machine learning-based system for different leak sizes was developed and to train the machine learning models, various statistical metrics were retrieved from the AE signal with a sliding window technique using an adaptive threshold to preserve the characteristics of different emissions [8]. For each AE sensor data category, three AE sensor datasets were collected and 11-time and 14-frequency domain features for a one-second window were extracted. Feature vectors were created using the measurements and the information that went along with them. Using the four datasets about gas and water leaks at varying pressures and pinhole leak sizes and several well-known classifiers, a remarkable total classification accuracy of 99% was attained, yielding dependable and efficient outcomes appropriate for the deployment [8]. Long Short-Term Memory Autoencoder was demonstrated to have the ability to identify the operational condition of the pipelines under monitoring [9]. Further studies using the support vector machine (SVM) and the decision tree algorithm (DT) to identify leaks in water pipes with two operational parameters; the flow rate and flow, at both the inlet and outlet of an experimental pipe were done and a leak detection technique with an accuracy of 97% was developed [10].

Though there exist some differences between the experimental data and the actual pipeline operation data, Particle Swarm Optimization Algorithm Optimized SVM highly suitable for pipeline leak identification [11]. An alternative refrigerant gas leak detection system is presented to reduce human intervention and boost confidence in the system in case of leakage and by using experimentation cycles including factors of the industrial environment, product, and a thermographic camera with infrared technology to get the dataset, regression logistic method performed the best in predictions, demonstrating that

Vol. 6, No. 1, March 2024

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

automatic decision-making in an industrial environment concerning gas leaks [12]. Intelligent models are robust and have reduced the average time to leak detection by a tangible percentage for all the intelligent models compared to a real-time transient model in literature and all had a time savings of 25% to 48% [13].

Information and Communication Technology (ICT) has been leveraged towards the attainment of environmental sustainability globally as diverse environmental challenges such as pollution and loss of biodiversity [14]. Having examined some relevant literature related to this research, some research gaps exist, including manual, partially automated, and/or inconsistent detections of leaks, which this research intends to bridge.

2. METHODS

Sensing devices to monitor the external part of the pipelines helps in determining abnormalities in the pipeline and this may require physical contact using such devices as acoustic sensing, and fibre optic sensing [15]. A system founded on the fundamental principles of mass, momentum, and energy conservation is used. This technique is based on the idea that a leak produces its transient signals, and the pressure and flow waves it creates travel to the pipeline's ends, leaving an imprint on the measured data and the flow, pressure, temperature, and density at specific receiving and delivery locations will often compute or assume an approximate fluid flow, pressure, and temperature for the whole pipeline [16].

The following presumptions will be utilized in this work to model natural gas flow through a pipeline. To create a model, certain assumptions must be established;

1) The cross-sectional area of a pipeline is constant.
2) There is a horizontal or near-horizontal flow.
3) The fluid is homogeneous, compressible, and in a single phase.
4) There are no chemical interactions happening between the fluid and the pipe.

The following are the governing equations:

$$Mass=\frac{\partial_{\varrho}}{\partial_{t}}+\frac{\partial_{\varrho u}}{\partial_{x}}=\{\} \tag{1}$$

$$Momentum= \frac{\partial}{\partial_{t}}(\varrho_{u}\partial_{x})+\frac{\partial_{\varrho u}}{\partial_{x}}(p+\varrho_{u^2})\partial_{x}=F_{g}+F_{f} \tag{2}$$

$$Energy=\frac{\partial}{\partial_{t}}T+C_{\varrho}u\frac{\partial}{\partial_{x}}T=q_{h}\text{-}g_{u}sin\theta \tag{3}$$

Vol. 6, No. 1, March 2024
p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

Adding Equations 1, 2, and 3:

$$\frac{\partial}{\partial_t}\left(p+\varrho_u+T_\varrho\right)+\frac{\partial}{\partial_x}\left(\varrho_u+p+\varrho_{u^2}+T\varrho_u\right)=\beta+\gamma \tag{4}$$

Where:

$$\beta=-\left(\varrho g sin\theta+\frac{\varrho u^{2f}}{2D}\right), \gamma=\frac{1}{c^1}\left(\varrho q_h-\varrho g sin\theta\right)\ q_h=\frac{4U_h(T_a-T)}{D} \tag{5}$$

Writing equation 5 in vector form

$$\frac{\vec{\partial}_v}{\partial_t}+\frac{\vec{\partial}_f}{\partial_x}+\vec{D}=0$$

Where

$$\vec{v}=\begin{pmatrix}\varrho\\ \varrho u\\ T\varrho\end{pmatrix};\vec{F}=\begin{pmatrix}\varrho u\\ P+\varrho u\\ T\varrho u\end{pmatrix};\vec{D}=\begin{pmatrix}0\\ \beta\\ \gamma\end{pmatrix}$$

Three different methods are often applied to solve equation 6. These methods include explicit Procedures, descriptive techniques, and Characteristics Methods. This intelligent model is data-driven and can be solved as approximations of equation 6 using data to fully characterize it. With these standards in mind, the Algoritms listed below were chosen: Artificial Neural Network (ANN), Support Vector Machine (SVM), Gradient Boosting (GB), Random Forest (RF), and Decision Tree (DT) and the dataset collected from kaggle.com is as in Figure 1 after dimensionality reduction. To reduce redundant information in the dataset, dimensionality reduction is needed [17]. To this end.

llhead Temp. (C)	Wellhead Press (psi)	MMCFD-gas	BOPD (barrel of oil p	BWPD (barrel of wate	BSW - basic solid an	CO2 mol. (%) @ 25 (	Gas Grav.	CR-corrosion (
64.13	2058.81	2.53	1307.94	5815.68	21.06	4.1099	0.7434	0
68.21	1883.68	2.73	610.06	6343.57	9.71	0.933	0.7421	0
45.27	948.74	3.9	480.06	6251.32	23.71	4.1899	0.7915	0
66.97	2036.34	15.26	700.38	7795.69	61.04	1.6463	0.9139	0
67.21	1340.54	10.36	1209.46	123.96	20.47	2.1917	0.8877	
46.44	939.73	0.96	1586.07	2441.85	70.71	3.1392	0.7428	0
64.93	1397.75	5.04	1639.18	3132.28	23.56	2.0979	0.8144	
55.51	1635.42	0.42	724.31	2371.9	66.82	3.6023	0.8231	0
66.63	1945.21	7.92	1117.32	8980.52	46.65	1.6905	0.7849	0
66.9	1537.04	11.81	1432.65	7300.19	16.33	2.0667	0.792	0
68.26	568.19	4.67	1381.38	585.66	23.17	2.0838	0.8788	0
73.06	977.02	10.17	1094.82	5639.4	58.5	2.0721	0.8352	0
56.27	1613.28	4.99	392.29	8568.67	71.69	3.8541	0.721	

Figure 1. Snapshot of the dataset

3. RESULTS AND DISCUSSION

The result obtained from the design; implementation of the oil pipeline detection system is presented in this section.

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

3.1. Model Training and Evaluation

The training of the model was done using cross validation technique. In cross validation, a portion of the training data is used for validation. The model was trained on the k-fold (10 folds) set aside as the training dataset. During the process of training the model, 1 out of the 10 folds was reserved for validation on each of the training instance. In each instance of the training, a validation set is used to validate the model performance. At the completion of the ten training instances, the performance of the each of the model instance was evaluated and the parameters of the best performing model was selected as the model for evaluation on the test dataset. Figure 2 shows cross validation process used in the training process.

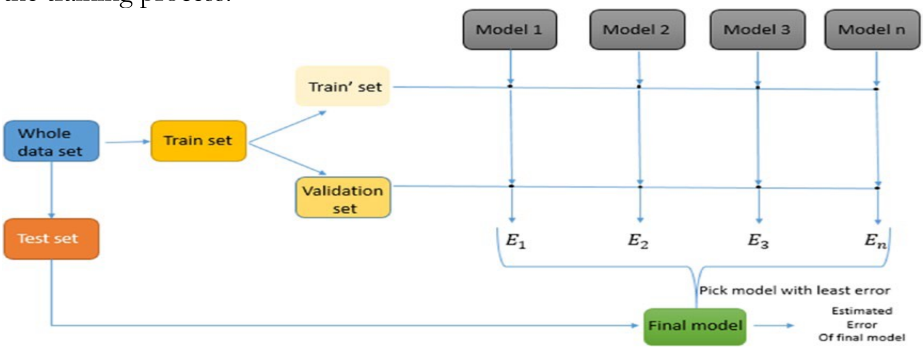

Figure 2. Cross validation training process

3.2. Hyperparameter Tuning

Hyperparameters are used in order to determine and build an optimal model. In the process of analysing the performance of a model, we used the cross validation to search through different parameter space for the most efficient and optimal values of the project algorithm's hyperparameters. There strategies that are frequently used for the determination of the most optimal hyperparameters and this include the grid search, random search and Bayesian optimization. Though hyperparameter optimization is frequently done using grid search since it is parallelizable and straightforward but it takes on greater significance when working with a big dataset with the inability to automatically set the grid hence adding a human aspect to a procedure that would be better served by an entirely automated system.

The conventional grid search has been modified by randomized parameter search which accepts input for grid components and input for distributions. Considering the parameter gamma, the values of which were provided directly when using grid search. In the random search, a distribution from which gamma was to be sampled was provided. The effectiveness of randomized parameter

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

search is based on the fact that the hyper-parameter optimization functions often have low dimensionality and that certain factors have a greater impact than others. By indicating the number of iterations the intended frequency for random parameter sampling is achieved. A snapshot of the code showing hyper-parameter tuning with random search is shown in Figure 3. The result of the hyper-parameter tuning shown in Figure 3 highlights the best parameter obtained from the search.

```
# Creating model pipeline
pipe = Pipeline([("scaler", preprocessing.StandardScaler()),
                 ("Classifier", SVC(random_state = seed))])
# Searching parameters
params = [{"Classifier__kernel": ["rbf", "linear"],
           "Classifier__gamma": [1, 0.1, 0.01, 0.001],
           "Classifier__C": [1, 10, 100, 1000]}
         ]
# Creating grid
svc_clf_grid = RandomizedSearchCV(estimator = pipe,
                                  param_distributions = params,
                                  cv = StratifiedKFold(n_splits = 10,
                                                       shuffle = True,
                                                       random_state = seed),
                                  n_iter = 10,
                                  verbose = 2,
                                  scoring = "accuracy",
                                  n_jobs = -1)
# Fit the model
svc_model = svc_clf_grid.fit(X_trainStandart, y_train)

# Get best parameters
print("Best parameters for SVC model: ", svc_model.best_params_)
```

```
Fitting 10 folds for each of 10 candidates, totalling 100 fits
Best parameters for SVC model:  {'Classifier__kernel': 'rbf', 'Classifier__gamma': 0.01, 'Classifier__C': 1000}
```

Figure 3. Hyper-parameter tuning

3.3. Model Performance Evaluation Metrics

When evaluating the performance of a machine learning model, several metrics can be used depending on the type of problem being solved and the desired outcome. This study is based on the detection of pipeline leakages. Typical metrics used for evaluation are used in this work and are explained as follows:

1) **Accuracy:** This represents the number of correct predictions made by the model divided by the total number of predictions expressed Equation 7.

$$Accuracy=\frac{T_p+T_n}{T_p+T_n+F_p+F_n} \quad (7)$$

2) **Precision and recall:** These metrics are used when the goal is to minimize false positives or false negatives and represent the ratio of true positives to the total number of predicted positives, while recall is the ratio of true positives to the total number of actual positives. The mathematical representation of the precision and recall are expressed in Equation 8 and Equation 9.

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

$$Precision=\frac{T_P}{T_P+F_p} \tag{8}$$

$$Recall=\frac{T_p}{T_n+F_p} \tag{9}$$

3) **F1 score:** This is a weighted average of precision and recall and is commonly used when the goal is to balance the trade-off between precision and recall. It is expressed in Equation 10.

$$F1_{score}=2\times\frac{precision \times recall}{precision+recall} \tag{10}$$

4) **Confusion Matrix:** The number of precise and unreliable predictions is categorized in a table called the confusion matrix. It summarizes the effectiveness of a classification model when applied to a set of test data for which the true values are known. The confusion matrix reveals how many relevant samples the classifier has identified and how many of them have been accurately identified. This recall ratio provides a measurement of inclusiveness.

3.3. Model Results Evaluation

ML algorithms are highly efficient in clinical decision support systems and help in timely interventions for enhanced health care delivery [18]. This is confirmed in the results obtained from evaluating the model as presented in this section. The result of the classification report is presented in Table 1. The summary of the classification report contains the scores for accuracy, precision, F1-Score, and recall. The classification report shown in Table 1 shows how well the model performed when measured against established standards for evaluating machine learning classification problems. The test dataset, which was separated and reserved from the original data, was used to assess the model performance. The result of the evaluation using the test data showed that the model achieved an accuracy of 97%. On the analysis of each class, 97% precision, recall, and F1 scores were obtained. The result obtained from the classification report shows that the model performed well in the detection and classification of oil pipeline leakages.

The confusion matrix obtained during the evaluation of the project model is presented in Figure 3 and the result shows the comparison between the predicted classes and true values. In the result, a total of 3088 observations were used as test data points to assess the model performance. Low and high are represented by 0, and 1 in the confusion matrix. Out of the total sample of 3088,

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

1436 samples constituted low out of which 1,393 were correctly classified while 43 were misclassified as high. On the other hand, 1,652 cases belong to high out of which 1,608 were correctly classified while 44 were misclassified as low. The result of the confusion matrix aligns with the classification report which further buttresses the fact that the model performed well in classifying the two classes of chances of pipeline leakage occurring due to corrosion defect based on the well and sensors parameter values.

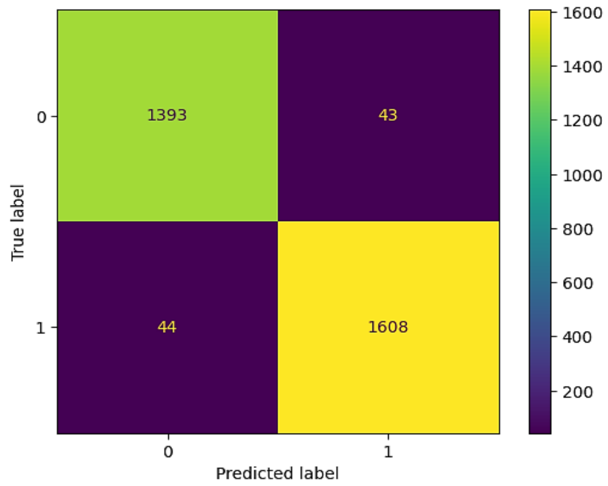

Figure. 4. Confusion matrix

The model prediction rate representing the percentage of correct prediction of each class, a purely qualitative measure of model performance shows the prediction success rate for chances of occurrence of oil leakage. A complete comparison of the performance metrics used in assessing the performance of the model using train and test dataset is presented in Table 2.

3.4. Model Prediction Rate

The prediction rate of the model must be included, even if there have been discussions and implementations of more comprehensible and acceptable metrics for the evaluation of categorization problems. The result obtained from the prediction rate which is expressed in percentages shows that 97% was obtained as the success rate of the SVM model when predicting on the test data. This finding suggests that the support vector machine learning model, which was created and implemented to predict oil leaks, accurately identified 97% of the test data.

Journal of Information Systems and Informatics
Vol. 6, No. 1, March 2024
p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

3.5. Model Performance Metrics Comparison

Figure 5 shows the comparison of the different classification model evaluation metrics applied to the training and testing data. As expected, the results obtained for accuracy, precision, recall, and F1 score were higher than scores obtained when the model was evaluated using test data. This discrepancy is easily explained due to the fact training data are seen by the model. Nonetheless, the derivation from the metrics comparison however shows that the model performed well when inferences were carried out on the test data. Identifying the appropriate coefficients that match the optimal value of the cost function is known as model fitting [17]. Here, the differences in the result of the evaluation of the model using test and train data indicate that the model did not overfit. Moreover, the result also implies that the training and testing errors of the model assessment are quite low. Figure 6, Figure 7, and Figure 8 show the box plot of Corrosion defect per each feature variable used while Figure 9 shows the Box Plot of Feature Variables.

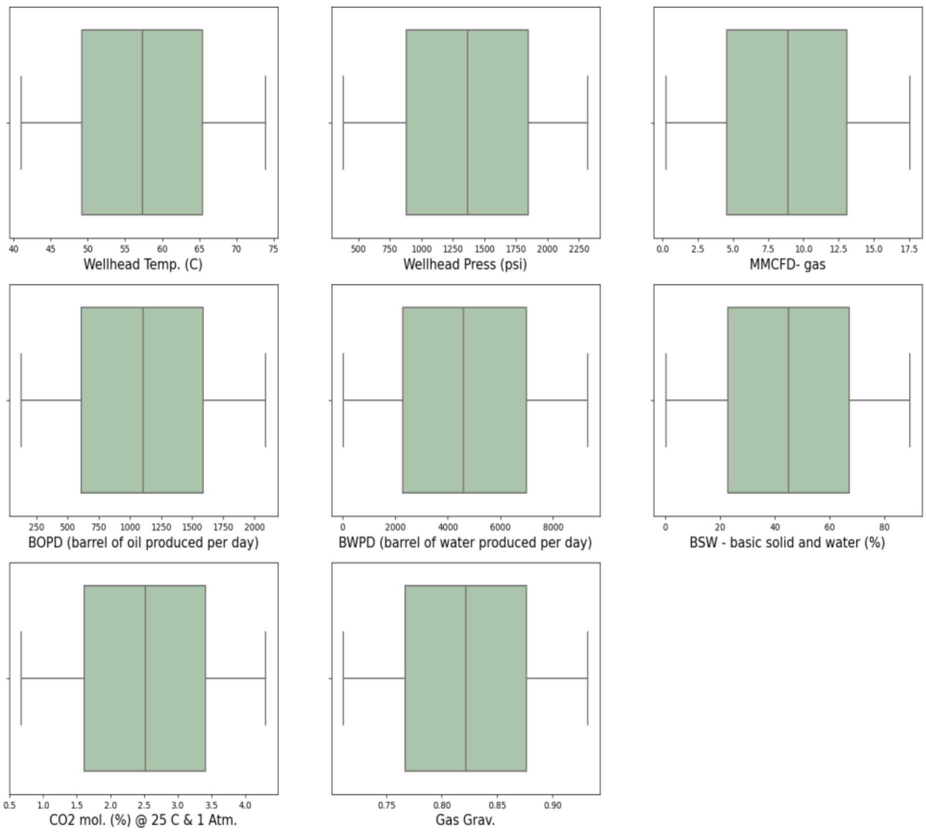

Figure 5. Box Plot of Corrosion defect per each feature variable

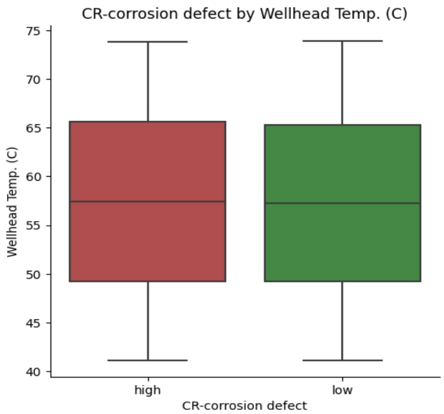

Figure 6. Corrosion defect vs Wellhead Pressure

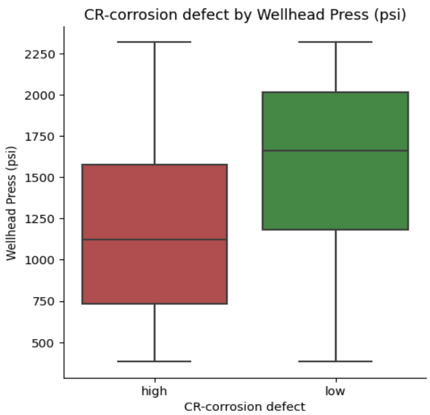

Figure 7. Corrosion defect vs BSW

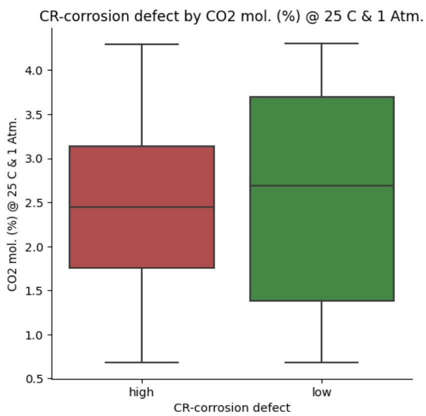

Figure 8: Comparison of the different classification model evaluation metrics

p-ISSN: 2656-5935 http://journal-isi.org/index.php/isi e-ISSN: 2656-4882

Data visualization enables us to analyze data when we are unsure of the precise questions, we need to ask beforehand. Figure 10 follows data from the scatter plot of feature variables.

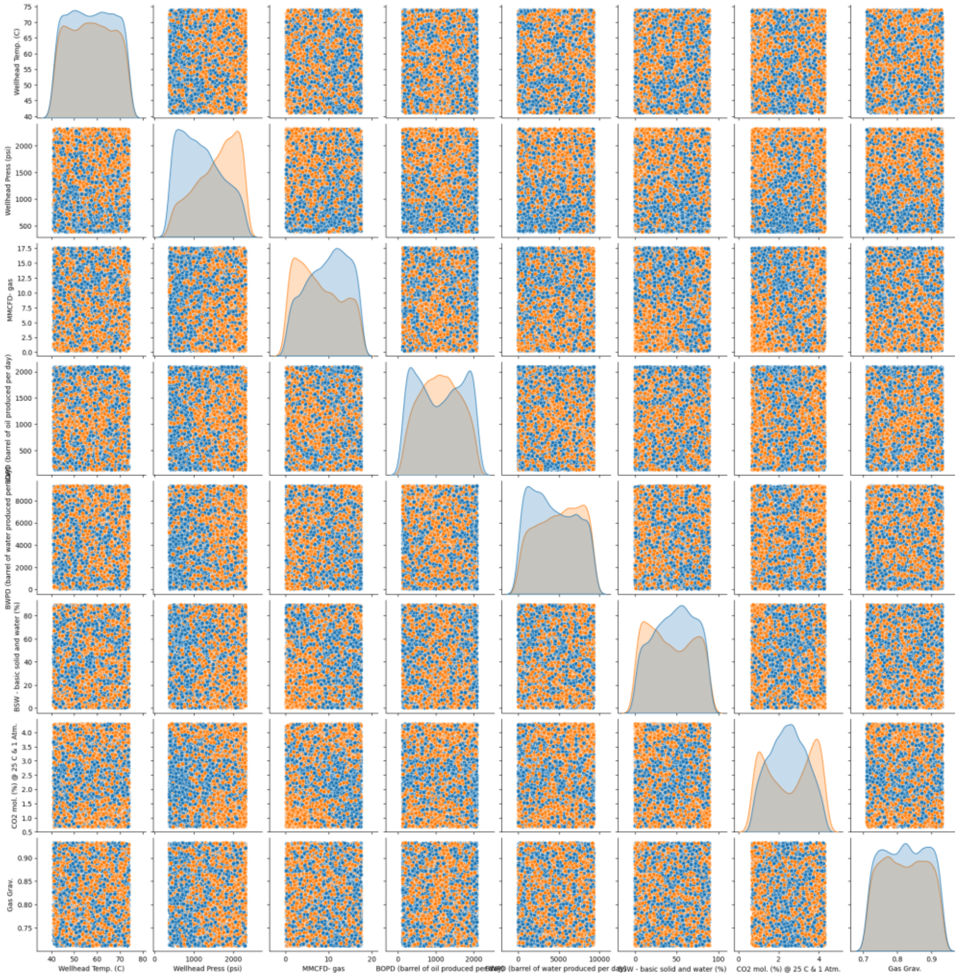

Figure 9. Box Plot of Feature Variables

4. CONCLUSION

Considering the grave implications of leakage to the environment and the organizations, this study embarked on exploiting the capabilities of machine learning algorithms in solving these problems. Having collected a comprehensive dataset on oil and pipeline leakage. Preprocessing data cleaning, data normalization, feature analysis, and data feature extraction were carried out on the dataset and the cleaned data was trained with the CNN model developed, the trained model was integrated into the monitoring system to provide notifications of leaks with an accuracy of 97%. The model performed better than the existing systems. This demonstrates the increasing importance of machine learning in the energy sector and is better suited for deployment in various pipeline networks with the potential to contribute to the overall improvement of pipeline safety and sustainability.

REFERENCES

[1] C. Lirong, W. Dingxiong, W. Hui, W. Xiaodong, G. Landing, M. Fen, X. Fei, F. Muqun, Y. Zhaofeng, L. Chao, J. Tiancheng, "Research on an oil pipeline anomaly identification method for distinguishing true and false anomalies," *Mobile Information Systems*, vol. 2022.

[2] O. Muhammad, H. Ali, "Deep Learning approach for oil pipeline leakage detection using image-based edge detection techniques," *Journal Européen des Systèmes Automatisés*, vol. 56, no. 4, pp. 663-673, 2023.

[3] W. Dongmei, S. Shaoxiong, L. Jingyi, H. Zhongrui & C. Jing, "Research on gas pipeline leakage model identification driven by digital twin," *Systems Science & Control Engineering*, vol. 11, no. 1, 2023.

[4] J. Renato, O. Thiago; P. Martin, "Leakage prevention and real-time internal detection in pipelines using a built-in wireless information and communication network." *SPE journal,* vol. 25, no. 5, pp. 2496–2507, 2020.

[5] A. Ekong, B. Ekong, A. Edet, "Supervised machine learning model for effective classification of patients with covid-19 symptoms based on bayesian belief network," *Journal of Information Systems and Informatics*," vol. 2, no. 1, 27-33, 2022.

[6] A. Ekong, A. Etuk., S. Inyang, M. Ekere-obong, "Securing against zero-day attacks: a machine learning approach for classification and organizations' perception of its impact," *Journal of Information Systems and Informatics,* vol. 5, no. 3, pp. 1123-1140, 2023.

[7] J. Choi, S. Im, "Application of CNN models to detect and classify leakages in water pipelines using magnitude spectra of vibration sound," *Applied Sciences, vol. 13. no.* 5, 2023.

[8] N. Ullah, Z. Ahmed, J. Kim "Pipeline leakage detection using acoustic emission and machine learning algorithms," *Sensors (basel)*, vol. 23, no. 6 pp. 3226, 2023.

[9] S. Christos, T. Panayiotis, G. Fotis, G. Nektarios, P. Aret, "Evaluation of deep learning approaches for oil & gas pipeline leak detection using wireless sensor networks," *Engineering Applications of Artificial Intelligence,* vol. 113, 2022.

p-ISSN: **2656-5935** http://journal-isi.org/index.php/isi e-ISSN: **2656-4882**

[10] A. July and C. Gomez, "Application of machine learning techniques for leak detection in horizontal pipelines" *Webology*, vol. 19, no. 4, 2022.
[11] G. Jianfeng, Z. Yu, N. Kai, Z. Huaizhi, H. Bin, Y. Jialin, "Research on oil-gas Pipeline Leakage Detection Method Based on Particle Swarm Optimization Algorithm Optimized Support Vector Machine," *Journal of Physics: Conference Series*, vol. 2076, 2021.
[12] S. Moreira, V. Shah, M. Varela, A. Monteiro, G. Putnik "Supervised machine learning applied to gas leak detection in air conditioner cooling system," *The 14th international conference on axiomatic design iop conf. series*," *Materials Science and Engineering,* no. pp. 1174 012008, 2021.
[13] A. Oluwatoyin, O. Adebayo, "Leak detection in natural gas pipelines using intelligent models," *SPE Nigeria annual international conference and exhibition*, no 198738, 2019.
[14[C. Owuru, J. Osakwe, S. Akinsola, "Utilisation of Information and Communication Technology for Environmental Sustainability: A Global Perspective" *Journal of Information Systems and Informatics*, vol. 3, no 3, pp. 403 -423, 2021.
[15] A. Mutiu, F. Wai-Keung, K. Aditya, "Recent advances in pipeline monitoring and oil leakage detection technologies: principles and approaches," *Sensors,* vol. 19, no. 11, pp. 1-36, 2019.
[16] R. Balda, K. Civan, W. Companies, "Application of mass balance and transient flow modeling for leak detection in liquid pipelines," *SPE Oklahoma: Society of Petroleum Engineers*, no. 164520, 2013.
[17] A. Ekong A., "Evaluation of machine learning techniques towards early detection of cardiovascular diseases," *American Journal of Artificial Intelligence*, vol. 7, no. 1, pp. 6-16, 2023.
[18] A. Ekong, G. Ansa, N. Odikwa, E. Adabra, "A hybrid machine learning model for classifying phishing uniform resource locators," *International Journal of Engineering Research in Computer Science and Engineering*, vol. 10, no. 5, pp. 63-67, 2023.
[19] B. Ekong, O. Ekong, A. Silas, A. Edet, B. William, 'Machine Learning Approach for Classification of Sickle Cell Anemia in Teenagers Based on Bayesian Network," *Journal of Information Systems and Informatics*, vol. 5, no. 4, pp.1793-1808, 2023.