MULTIDISCIPLINARY SCIENCE JOURNAL

RESEARCH ARTICLE
Published Online: September 26, 2024
https://doi.org/10.31893/multiscience.2025137

Implementation of machine learning methods in predicting failures in electrical submersible pump machines

Singgih Saptadi[a] ✉ | Achmad Widodo[b] | Muhammad Fitrah Athaillah[a] | Muhammad Farid Ayyasyi[a]

[a]Industrial Engineering and Management, Department of Industrial Engineering, Faculty of Engineering, Diponegoro University, Indonesia.
[b]Mechanical Engineering, Department of Mechanical Engineering, Faculty of Engineering, Diponegoro University, Indonesia laboratory, Indonesia.

Abstract The Electric Submersible Pump (ESP) is one of the artificial lift tools widely used in oil and gas wells. Approximately 25% of oil and gas wells have been equipped with ESP. The ESP unit excels in lifting fluid rates much larger than most other types of artificial lift. The purpose of this research is to develop a machine learning model that can predict the likelihood of ESP installation failure. By making predictions at the early stage, necessary actions can be prepared to prevent and anticipate these issues. Thus, the reduction in the amount of produced fluid can be minimized. The model in this study uses 8 input parameters, namely, motor Ampere, Frequency, Pump Intake Pressure, Temperature Motor, Output Volt, Pump Discharge Pressure, Input Voltage, and Motor Horse Power (HP). There are also 8 failure classifications for which the model will be created. The initial stage involves collecting and cleaning raw data and processing it to detect anomalies. Once the data is processed, it proceeds to the model formation and model validation stage. The method used for the model is based on the comparison of two methods, namely, decision tree and k-nearest neighbor. The output of the research is a dashboard that can display visualizations of data and the prediction results from the created model. Millions of data points are used to create a database consisting of 77 wells over the past 2 years. The model developed generates several types of failures that can be read. These failures have their own characteristic parameters. The model obtained has an accuracy rate above 90%. The output of this model includes the most likely failure prediction based on the latest input data, thus effectively addressing well problems, warning of impending failures, reducing failures, and assisting in scheduling ESP repairs and maintenance.

Keywords: ESP, machine learning, failure prediction, decision tree, K-nearest neighbor

1. Introduction

As time progresses, the age of the produced well experiences a decline in pressure over time. There is a point where the natural reservoir pressure is no e allonger sufficient to flow and lift reservoir fluids to the surface. The electric submersible pump (ESP) is an artificial lift tool widely used in oil and gas wells (Panbarasan et al., 2022). It is estimated that 25% of oil and gas wells have been equipped with ESPs (Adesanwo et al., 2016). Compared with other artificial lift systems, the main advantage of electrical submersible pumps their high production capacity (Abdelaziz et al., 2017).

The ESP system consists of several interacting elements and is fundamentally complex. The behavior of the ESP system varies significantly due to differences in fluid settings, reservoir conditions, well completion methods, and various types of equipment involved (Grasiian et al., 2017). The installation of ESPs can lead to increased production costs. However, there is also an expectation of economic benefits from the implementation of the ESP because it can increase or maintain the level of oil production. An ESP (electrical submersible pump) is also a tool that can withstand extreme environments that are physically inaccessible to operation and maintenance teams. These extreme conditions can increase the likelihood of ESP failure and have a negative effect on the lifespan of ESP equipment, as well as a decrease in company revenue (Fang et al., 2021).

The ESP itself is equipped with downhole monitoring sensors that transmit data, such as motor temperature, pump inlet pressure, motor vibration temperature, and motor current, to the surface. These data can be used in the development of key failure condition indicators for ESP machines. For example, if the motor temperature increases drastically, the pump will shut down automatically. Other indicators may include changes in pressure, current, and other parameters. The data from these sensors are collected and managed in a database. However, the database system is designed only to store historical data over time and is not designed to analyze data, generate reports, or provide real-time intervention.

Owing to its use, the ESP machine will gradually experience deterioration in its condition, which can lead to failure. In this case, predictive maintenance is a solution that can be used to reduce these failures by monitoring trends in installed sensor data to predict failures in ESP machines (Sharma et al., 2022; Silvia and Furlong, 2023; Melo et al., 2023; Bermudez et al., 2021). The development of a specifically tailored failure prediction model for ESP machines provides benefits for drilling operations.

 Received: July 23, 2024 | Accepted: August 24, 2024

This is especially true in monitoring and evaluation, as well as in testing prescriptive models that can be used to improve decision-making and operational management of ESPs (Lastra and Xiao, 2022). Failure prediction in ESP machines is a crucial technical requirement in drilling operations and has the potential to be applied to various oil and gas industry equipment, thereby preserving machine life and enhancing oil production (Grasiian et al., 2017).

Failures of electric submersible pump (ESP) machines are common in the oil industry. On the basis of data from the company under study, there were a total of 1860 failures within a period of 2 years. Late problem identification can result in the loss of fluid production and potential damage to the ESP system itself. Therefore, there is a need for a system that can detect ESP failure issues, read sensor data in real time, and determine the problems corresponding to the sensor data. A similar concept was established by Bermudez et al. (2014) and Adesanwo et al. (2016) in implementing artificial intelligence called fuzzy logic, which was further refined by Grasiian et al. (2017). In this paper, an improvement is made by applying machine learning to analyze the data. Machine learning is designed to create a model that becomes more accurate as more data are added. Machine learning can detect problems by building a model that represents the dataset used (Nasteski, 2017).

On the basis of the issues mentioned above, research has focused on two aspects: predicting failures in ESP machines and predicting the remaining operational time before ESP machine failures occur. Through this research, it is hoped that an accurate predictive model can be created, leading to more precise maintenance of ESP machines before failures or malfunctions occur, thereby minimizing the financial losses incurred by the company.

1.1. Research Objectives

1. Potential issues for each ESP machine failure are identified on the basis of historical well data in the company's database via machine learning.
2. To develop a machine learning model that can predict future failures to assist in decision-making and management within the ESP system.

2. Materials and Methods

2.1. Data collection

The data originate from a well within the company's working area, which is equipped with various preinstalled sensors. The raw data obtained have different time intervals and need to be prepared before processing. The raw well data obtained are from a separate company for each month. Therefore, all the raw data for each month are combined, starting from October 2020 to August 2022. The total overall dataset consists of 10,089,493 records and has 28 columns, and all the data types are still in object format. Furthermore, there are LPO data containing historical damage labels that have been recorded to have occurred on the ESP machine within a specified period. The LPO data consists of 28 columns with 1860 recorded damages, both planned and unplanned.

The raw data need to be filtered before processing. The data are filtered by removing unnecessary variables, changing the data types as needed, and eliminating missing values in the available data. The goal is to produce more accurate data for modeling. In this study, several sensor data columns, including the Motor Ampere, Frequency Motor, Pump Intake Pressure, Temperature, OutputVolt, Pump Discharge Pressure, Input Voltage, and Motor Horse Power (HP), which can be seen in Figure 1 and Figure 2, are utilized as indicators for the ESP machine. For label data, unplanned types of damage or those occurring suddenly are used. Therefore, it is necessary to remove some columns that are not used to facilitate data processing.

The failure modes used in this study can be seen in Table 1, which include normal, mechanical, electrical, control and system failure modes, and thermal failure modes. These types of failures are the output results of the machine learning prediction.

2.2. Exploratory Data Analysis

Data exploration aims to analyze the data patterns that occur during failures. For each failure, an analysis is conducted on the patterns formed in the data, and the results are recorded. The same method of exploratory data analysis techniques was also applied by Bermudez et al. (Bermudez et al., 2014) and Adesanwo et al. (Adesanwo et al., 2016). Before the modeling stage, a classification of failures is conducted to categorize the different types of failures present in the data. Table 2 below explains the EDA results that occur on the sensor every time a failure occurs. C indicates that the data are constant.

2.3. Data Transformation

Data transformation is one of the fundamental steps in data preprocessing. When studying feature scaling techniques, we often encounter terms such as scaling, standardization, and normalization. The most commonly used technique in feature scaling is normalization. In machine learning and data mining, this process helps transform the values of numeric columns in a dataset to have a consistent scale. Normalization is frequently applied to data with different ranges or scales. There are several

commonly used normalization methods, such as min–max normalization, z score normalization, and minimal scaling normalization (Singh et al., 2015).

Standardization

$$x_{stand} = \frac{x - mean(x)}{standard\ deviation\ (x)} \quad (1)$$

Normalization

$$x_{stand} = \frac{x - min(x)}{max\ (x) - \min\ (x)} \quad (2)$$

Table 1 Failure mode.

Code	Failure Mode	Category
0	Normal	Normal
1	Overload	Mechanical Failures
	Underload	
	MDHP	
	Breaker Trip	
	Current Unbalance	
	Low Supply	
	High Supply	
2	EDHP	Electrical Failures
	Overcurrent	
	High Frequency	
	High Volt	
	Low Voltage	
	Undervolt	
3	Problem VSD	Control and System Failures
	Problem Panel	
	Unicom Controller Blank	
	UDMIN	
	ARK LOOK	
	Problem PCB Card	
	Problem PMSN	
4	Motor High Temp	Thermal Failures

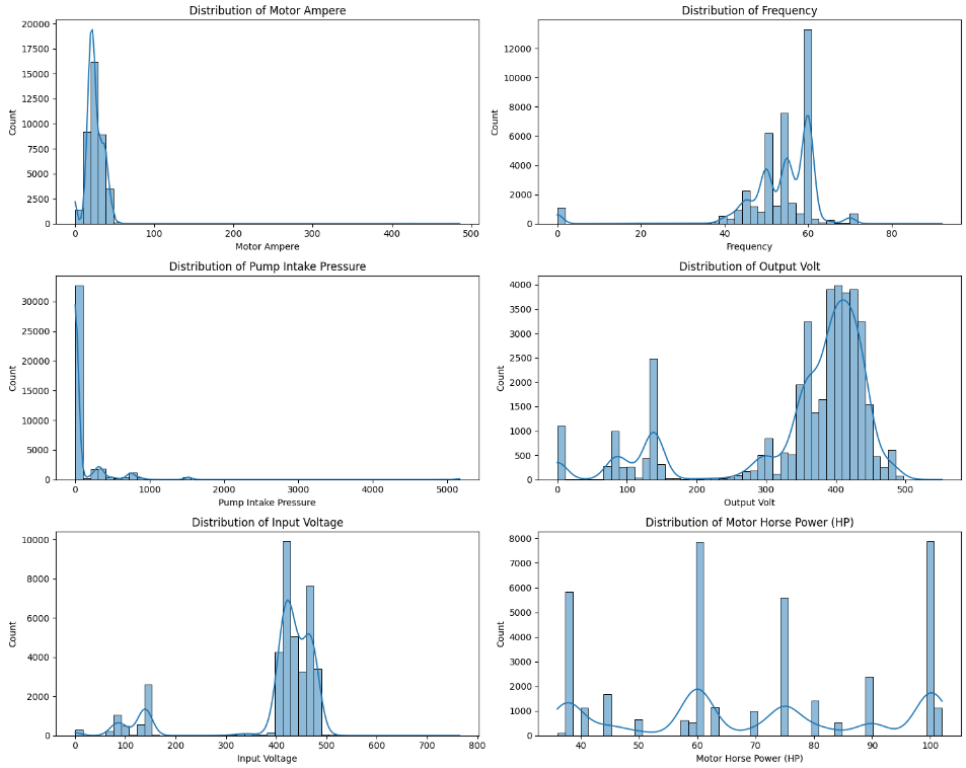

Figure 1 Data distribution.

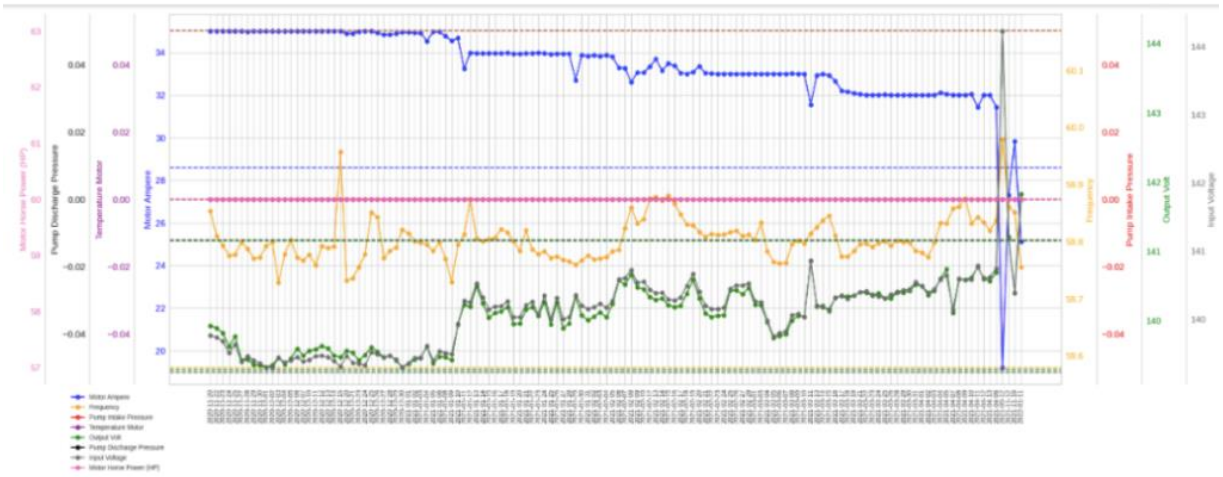

Figure 2 Data monitoring.

Table 2 Exploratory Data Analysis.

	Motor Ampere	Frequency	Pump Intake Pressure	Temperature Motor	Output Volt	Pump Discharge Pressure	Input Voltage	Motor Horse Power
Mechanical Failures	-	-	+	+	-	-	-	-
Electrical Failures	+	-	+	+	-	-	+	-
Control and System Failures	-	-	C	-	C	-	C	+
Thermal Failures	-	-	+	+	-	+	C	-

In this study, the raw data need to be filtered before processing. The data are filtered by removing unnecessary variables, changing the data types as needed, and eliminating missing values in the available data. The goal is to produce more accurate data for modeling. The data are then divided on the basis of the available well types to facilitate data analysis and simplify data labeling. Feature scaling is then performed to normalize the data so that the data ranges are not too distant from each other, which was also done by Grassian et al. (Grasiian et al., 2017). Figure 3 shows the data before transformation with normalization, while Figure 4 shows the data after normalization.

	Motor Ampere	Frequency	Pump Intake Pressure	Temperature Motor	Output Volt	Pump Discharge Pressure	Input Voltage	Motor Horse Power (HP)
0	20.978448	60.000000	764.700000	175.800000	442.275000	762.700000	420.670690	90.0
1	20.612281	58.947368	399.790526	91.909474	434.814737	398.744912	421.683860	90.0
2	20.997203	60.000000	0.000000	0.000000	442.801748	0.000000	421.136364	90.0
3	21.004643	60.000000	0.000000	0.000000	442.630000	0.000000	421.166786	90.0
4	21.022414	60.000000	0.000000	0.000000	442.639655	0.000000	420.982759	90.0

Figure 3 Data transformation before scaling.

	Motor Ampere	Frequency	Pump Intake Pressure	Temperature Motor	Output Volt	Pump Discharge Pressure	Input Voltage	Motor Horse Power (HP)
0	0.036251	0.278323	0.273937	0.269852	0.279102	0.269852	-0.155978	0.0
1	0.016038	0.201094	-1.734428	-1.764042	0.204307	-1.764042	-0.099172	0.0
2	0.037287	0.278323	-3.934769	-3.992352	0.284383	-3.992352	-0.129869	0.0
3	0.037697	0.278323	-3.934769	-3.992352	0.282661	-3.992352	-0.128163	0.0
4	0.038678	0.278323	-3.934769	-3.992352	0.282758	-3.992352	-0.138481	0.0

Figure 4 Data transformation after scaling.

2.4. Data processing and synthesis

The data processing process begins by labeling each well on the basis of the guidelines provided by the company. Then, outlier detection is carried out with the aim of removing data that lie outside the outlier. Data that are outside the outlier but detected as normal are eliminated. The goal is to improve accuracy during modeling. Before entering the modeling phase, the available raw data are consolidated into a dataset used for modeling. The next step is to categorize the types of faults into nine categories on the basis of the company's labeling guidelines. The obtained fault types include mechanical failures, electrical failures, control and system failures, and thermal failures. This is done to classify the existing faults for use in machine learning modeling. The prepared data then proceed to the modeling stage.

Owing to data imbalance, oversampling is needed. In this research, the oversampling method used was SMOTE.

The same method was also used by Zhen et al. (Zhen et al., 2023), who utilized SMOTE to overcome data imbalance. By using SMOTE, the minority class with fewer recorded data can be matched to the majority class with the most data. In this case, the majority class represents the normal operating condition where there are no issues in the well, so all categories of failures are excessively sampled to align with the quantity of normal data.

The steps of this algorithm are outlined below (Zhen et al., 2023):

Step 1: Randomly choose k points from the entire set of samples D = x1, x2, x3… xn and then define them as the sample cluster centers C1, C2, C3… Ck

Step 2: Calculate the distance between each sample and the specified sample cluster center. :

$$d = \sqrt{\sum(x_i - C_k)^2} \quad (3)$$

where x1,x2,x3, .,xi ∈ D; C1,C2,C3, .,CK ∈ C.

Step 3: assign the samples to the closest sample clusters:

$$x^i \in C_{nearest} \quad (4)$$

Step 4: Recalculate the new cluster centers for the samples.

$$\mu_i = \frac{1}{|C_I|}\sum_{x \in C_i} x \quad (5)$$

Step 5: Continue performing steps 2 through 4 until there are no further changes in any of the cluster centers.

Step 6: Generate additional minority samples by selecting clusters that contain either fewer or more minority classes and filtering them accordingly.

Step 7: Perform SMOTE oversampling for each filtered cluster of CK. :

$$X_{new} = x_c + rand(0,1) \times (\tilde{x} - x_c) \quad (6)$$

Xnew = new negative class sample generated above,

x_c = represents a random negative class selected from the m nearest neighbors within the filtered clusters,

$\tilde{x}$ = refers to the negative samples in the filtered clusters except for the m neighbors.

2.5. Modeling and Validation

2.5.1. Modeling

- *K*-nearest neighbor

K-nearest neighbor (KNN) represents a conventional technique in machine learning that has been adapted for extensive data mining endeavors. The fundamental concept involves employing a substantial training dataset, with each data point defined by a collection of attributes. Essentially, every point is positioned within a multidimensional framework, where each dimension represents a distinct variable. When we have a new data point (test), we want to find the K nearest neighbors that are closest to it (i.e., most "similar"). The value of K is usually chosen as the square root of N, the total number of points in the training dataset (Suyal and Goyal, 2022).

- *Decision Tree*

The decision tree is one of the models used in supervised learning in the field of machine learning. The decision tree model is commonly used to solve regression and classification problems, but it is more commonly used for classification tasks. Through this analysis, the aim is to anticipate the result on the basis of numerous input factors by continuously segmenting each factor into various potential results. The decision tree adopts a hierarchical arrangement comprising a root node, decision nodes, and leaf nodes. Leaf nodes represent the ultimate points that are not subject to further division and contribute to shaping the predictive outcomes of the decision tree (Nadiah et al., 2022).

2.5.2. Validation

Making decisions about appropriate evaluation metrics to assess the performance of an algorithm is a crucial step in this research. This is because classification models trained on imbalanced datasets may achieve high accuracy but tend to be biased toward the majority class. In this context, the classification report and confusion matrix are essential tools for evaluating classification models. n the confusion matrix, four terms are used to represent the results of the classification process. These terms are true positive (TP), true negative (TN), false positive (FP), and false negative (FN). The value of true negative (TN) indicates the number of negative samples correctly identified, whereas false positive (FP) indicates the number of negative samples incorrectly identified as positive. (Karsito and Susanti, 2019).

In this study, we also use evaluation metrics, namely, accuracy, precision, sensitivity, and the F1 score, and support the following equations (Erlin et al., 2022):

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (4)$$

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

$$Recall/Sensitivity = \frac{TP}{TP+FN} \quad (6)$$

$$Specificity = \frac{TN}{TN+FP} \quad (7)$$

$$F1 - Score = \frac{2\,(Recall \times precision)}{Recall + precision} \quad (8)$$

Precision: This metric measures the accuracy of the model in predicting positive classes.

Recall (Sensitivity): This metric measures the model's ability to detect all positive samples.

F1-score: This score combines precision and recall into a single metric that provides a comprehensive view of model performance.

Support: The number of actual samples for each class in the dataset.

3. Results

3.1. Modeling

The machine learning method was chosen on the basis of a comparison of several methods, namely, the K-nearest neighbor and decision tree methods. The data are divided into two sets: 80% for the training data and 20% for the test data. Even though the test data constitute only 20% of the data, the model created already has samples from all categories, meaning that the model can recognize all existing categories (Meddaoui et al., 2024). The input data include various features, such as Motor Ampere, Frequency Motor, Pump Intake Pressure, Temperature, OutputVolt, Pump Discharge Pressure, Input Voltage, and Motor Horse Power (HP). The output data include failure categories: normal, mechanical, electrical, control and system failures, and thermal failures.

On the basis of Figure 5 and Figure 6 above, both models (KNN and decision tree) without SMOTE exhibit a bias toward the majority class ('Normal'), with poor performance on minority classes. This is a common issue in imbalanced datasets. Applying SMOTE improves the models' ability to correctly classify minority classes. This is evident in the significant increase in correct predictions for 'Mechanical Failures', 'Electrical Failures', and other minority classes. After applying SMOTE, both the KNN and decision tree models achieve more balanced performance. However, specific nuances in the confusion matrices suggest that while both models benefit from SMOTE, the extent of improvement can vary on the basis of the model and the nature of the data. These results underscore the importance of addressing class imbalance to achieve a more accurate and reliable classification model in the context of predicting failures in electrical submersible pump machines.

It is evident that modeling without oversampling has higher accuracy than that with oversampling. This is due to the overfitting that occurs because of unbalanced data. To address this issue, oversampling via the SMOTE method is necessary (Tang et al., 2023). A validation comparison is then needed between the two methods via a confusion matrix. The results from the confusion matrix show that the best accuracy is achieved by the predictive model with oversampling, even though the displayed accuracy percentage is lower. In this case, modeling with oversampling is chosen because it provides better accuracy.

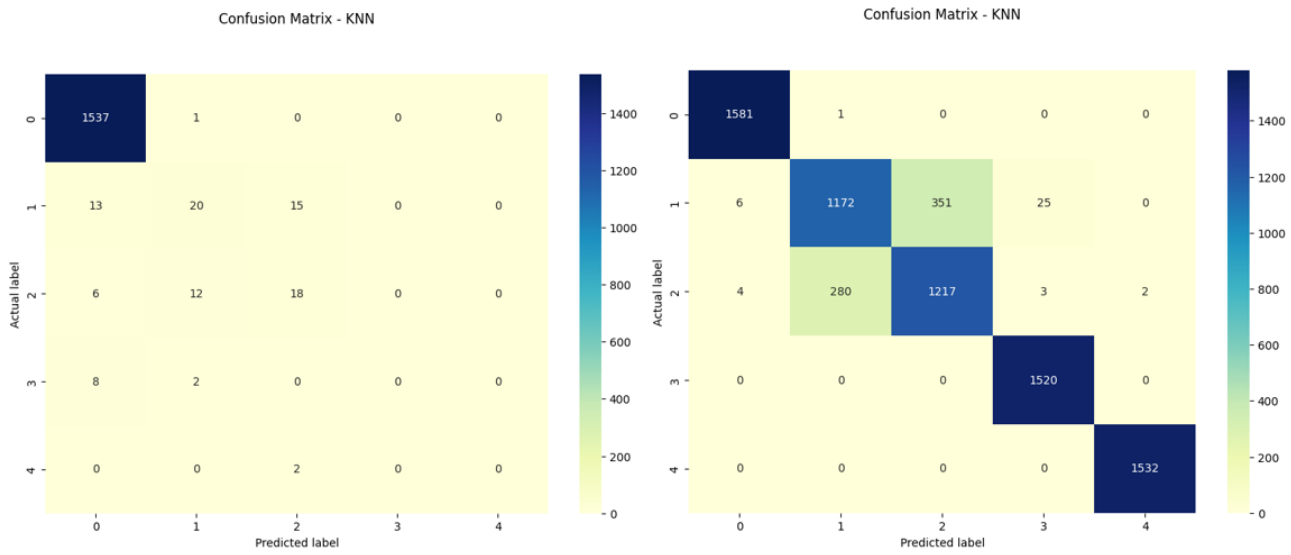

Figure 5 Confusion matrix of the k nearest neighbor before and after SMOTE.

Table 3 shows a summary of the models that have been created to predict the types of failures.

The Table 4 shows that there is no significant difference between the two methods used. Both methods have identical accuracy, precision, recall, and F1 score values. However, the K nearest neighbor has a higher accuracy of 0.9127 than does the

decision tree, which has an accuracy of 0.9108.

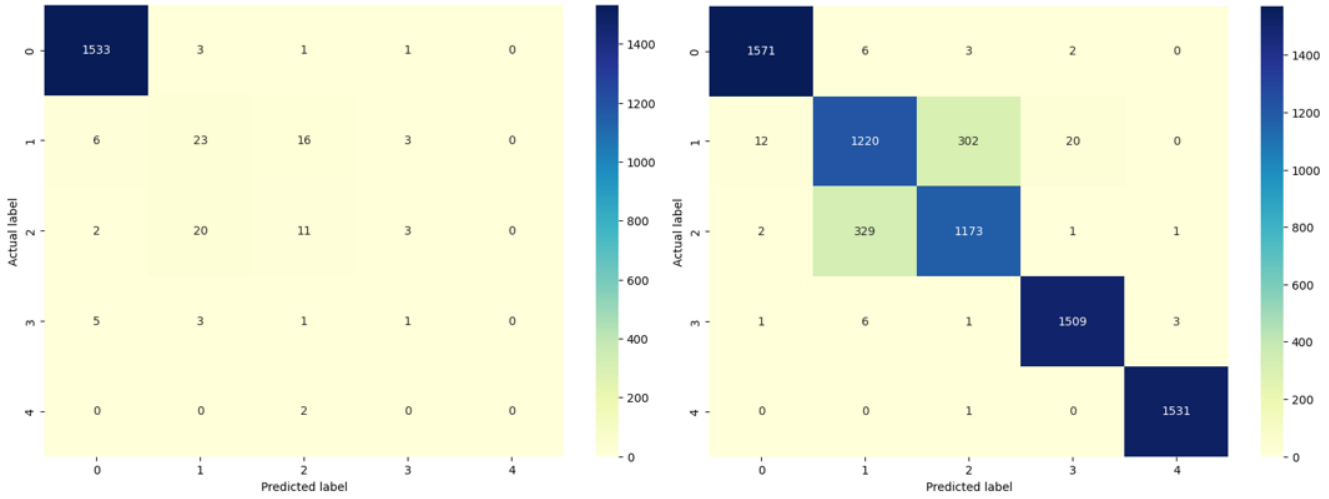

Figure 6 Confusion matrix of the decision tree before and after SMOTE.

3.2. Validation

The classification results from modeling with SMOTE are presented in Table 3.

Table 3 Classification report.

Model		Precision	Recall	F1-Score	Support
KNN with SMOTE	0	0.99	1	1	1582
	1	0.81	0.75	0.78	1554
	2	0.78	0.81	0.79	1506
	3	0.98	1	0.99	1520
	4	1	1	1	1532
Accuracy				0.91	7694
macro avg		0.91	0.91	0.91	7694
weighted avg		0.91	0.91	0.91	7694
Decision Tree with SMOTE	0	0.99	0.99	0.99	1582
	1	0.78	0.79	0.78	1554
	2	0.79	0.78	0.79	1506
	3	0.98	0.89	0.93	1520
	4	1	1	1	1532
Accuracy				0.91	7694
macro avg		0.91	0.91	0.91	7694
weighted avg		0.91	0.91	0.91	7694

Table 4 Accuracy report.

Method	Accuracy
K-Nearest Neighbor	0,9127
Decision Tree	0.9108

Overall, the application of SMOTE has resulted in significant improvements in the classification performance of both the KNN and decision tree models, especially for minority classes, demonstrating the importance of addressing class imbalance to achieve accurate and reliable classification models.

4. Discussion

The research on the implementation of machine learning methods in Electrical Submersible Pump (ESP) machines has identified multiple types of failures occurring over two years. These failures were grouped into categories: normal, mechanical failures, electrical failures, control and system failures, and thermal failures, which is also done in previous literature. (Abdalla et al., 2023; Mello et al., 2022). Failures often occur suddenly due to fluctuations in sensor indicators, as documented by previous studies on pump monitoring systems (Nguyen et al., 2020). Modeling was performed using machine learning techniques, specifically K-nearest neighbor (KNN) and decision tree algorithms, to predict future failures.

The implementation of these machine learning techniques for predicting failure types in ESP machines aligns with the

outcomes of previous research, demonstrating that KNN and decision tree algorithms are effective tools in predictive maintenance applications. (Salem et al., 2022). The accuracy of these models was significantly improved by implementing the synthetic minority over-sampling technique (SMOTE), consistent with results reported by Santoso et al. (2020), who demonstrated the effectiveness of SMOTE in handling imbalanced datasets in predictive models.

Observations from the confusion matrices reveal a substantial improvement in model performance after applying SMOTE, highlighting its critical role in enhancing accuracy and precision, particularly in minority class predictions. This finding aligns with the work of (Chawla et al., 2002), who originally developed SMOTE and showcased its capacity to address class imbalance issues in machine learning models. The visual representation provided by the confusion matrices underscores the impact of SMOTE, making it clear that the optimization achieved in this study is not merely coincidental but rooted in sound methodology and validated by previous research.

After testing, K-nearest neighbor emerged as the most optimal compared to decision trees, achieving an accuracy of 91.27% after applying SMOTE. This aligns with similar findings in predictive maintenance literature, where KNN has often been highlighted for its high performance in failure prediction tasks (Al-Ballam et al., 2023). This model's capacity to predict failure types with high precision offers a reliable tool for timely failure warnings, reducing the occurrence of failures and extending the operational life of ESP pumps. These results echo the broader consensus in the field that machine learning, particularly when coupled with class imbalance techniques like SMOTE, can significantly improve the reliability and functionality of predictive maintenance systems (Zhen et al., 2023).

This discussion effectively positions the study within the existing body of literature, emphasizing the innovative nature of the methodology and its alignment with prior research. The application of SMOTE not only enhanced model accuracy but also significantly improved predictive performance, particularly for minority failure classes. These findings highlight the critical role of addressing class imbalance in machine learning models, validating the methods employed in this study and underscoring its relevance to the field of predictive maintenance.

5. Conclusions

Research on the implementation of machine learning methods in electrical submersible pump machines has identified many types of failures that have occurred over a period of 2 years. After that, we grouped the failures according to the categories of damage, resulting in the identification of the following types of damage: normal, mechanical, electrical, control and system failures, and thermal failures. These failures occur suddenly on the basis of fluctuations in the sensor indicators. Modeling was then performed via machine learning to predict future failures.

The machine learning methods used to develop a model capable of predicting failure types in electrical submersible pump (ESP) machines are the k nearest neighbor (KNN) and decision tree methods. These models were compared in terms of their accuracy levels. The implementation of the SMOTE (synthetic minority oversampling technique) significantly increased the accuracy and precision of both machine learning algorithms. Observations from the confusion matrices reveal that the models provide more accurate results with SMOTE than those without SMOTE, demonstrating that SMOTE optimizes the accuracy of machine learning models. The confusion matrices provide a clear visual representation of the models' performance before and after applying SMOTE, with significant improvements in minority class predictions highlighting SMOTE's effectiveness in handling imbalanced datasets. These results underscore the importance of addressing class imbalance to achieve more accurate and reliable classification models for predicting failures in ESP machines.

The best algorithm model in this study is the K nearest neighbor. The classifier achieves an accuracy of 91.27% after the SMOTE procedure. The models can predict the type of failure that will occur with precise and functional accuracy, effectively addressing issues in the well. This enables timely failure warnings, reduces failures, and extends the life of ESP pumps.

Acknowledgments

The authors would like to thank the Faculty of Engineering, Diponegoro University, Indonesia, which provided financial support for this research. The authors also appreciate the help and cooperation of the participants who participated in this study.

Ethical considerations

Not applicable.

Conflict of interest

The authors declare no conflicts of interest.

Funding

This research did not receive any financial support.

References

Abdalla, R., Nikolaev, D., Gönzi, D., Manasipov, R., Schweiger, A., & Stundner, M. (2023). Deep Insight into Electrical Submersible Pump Maintenance: A Predictive Approach with Deep Learning. *SPE Offshore Europe Conference & Exhibition*.

Abdelaziz, M., Lastra, R., & Xiao, J. J. (2017). ESP Data Analytics: Predicting Failures for Improved Production Performance. *Abu Dhabi International Petroleum Exhibition & Conference*.

Adesanwo, M., Bello, O., Denney, T., & Lazarus, S. (2016). Prescriptive-Based Decision Support System for Online Real-Time Electrical Submersible Pump Operations Management. *SPE Intelligent Energy International Conference and Exhibition*.

Al-Ballam, S., Karami, H., & Devegowda, D. (2023). A Hybrid Physical and Machine Learning Model to Diagnose Failures in Electrical Submersible Pumps. *SPE/IADC Middle East Drilling Technology Conference and Exhibition.* Abu Dhabi.

Bermudez, F., Carvajal, G. A., Moricca, G., Dhar, J., Adam, F. M., Al-Jasmi, A., . . . Nasr, H. (2014). Fuzzy Logic Application to Monitor and Predict Unexpected Behavior in Electric Submersible Pumps (Part of the KwIDF Project). *SPE Intelligent Energy Conference & Exhibition*.

Bermudez, F., Nahhas, N. A., Yazdani, H., LeTan, M., & Shono, M. (2021). Unlocking the Potential of Electrical Submersible Pumps: The Successful Testing and Deployment of a Real-Time Artificially Intelligent System, for Failure Prediction, Run Life Extension, and Production Optimization . *Abu Dhabi International Petroleum Exhibition & Conference*.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic Minority Over-sampling Technique. *Journal of Artificial Intelligence Research*, 321–357.

Erlin, Desnelita, Y., Nasution, N., Suryati, L., & Zoromi, F. (2022). Impact of SMOTE on Random Forest Classifier Performance based onImbalanced Data. *Jurnal Manajemen, Teknik Informatika, dan Rekayasa Komputer*.

Fang, M., Raveendran, R., & Huang, B. (2021). Real-time Performance Monitoring of Electrical Submersible Pumps in SAGD Process. *IFAC-PapersOnLine Volume 54, Issue 11*, 139-144.

Grasiian, D., Bahatem, M., Scott, T., & Olsen, D. (2017). Application of a Fuzzy Expert System to Analyze and Anticipate ESP Failure Modes. *Society of Petroleum Engineers*.

Karsito, & Susanti, S. (2019). Klasifikasi Kelayakan Peserta Pengajuan Kredit Rumah dengan Algoritma Naive Bayes di Perumahan Azzura Residencia. *Jurnal Teknologi Pelita Bangsa*.

Lastra, R. A., & Xiao, J. (2022). Machine Learning Engine for Real-Time ESP Failure Detection and Diagnostics. *SPE Middle East Artificial Lift Conference and Exhibition*.

Meddaoui, A., Hachmoud, A., & Hain, M. (2024). Advanced ML for predictive maintenance: a case study on remaining useful life prediction and reliability enhancement. *The International Journal of Advanced Manufacturing Technology*, 323–335.

Mello, L. H., Oliveira-Santos, T., Varejão, F. M., Ribeiro, M. P., & Rodrigues, A. L. (2022). Ensemble of metric learners for improving electrical submersible pump fault diagnosis. *Journal of Petroleum Science and Engineering*.

Melo, R. A., Worth, D. J., & Swaffield, S. (2023). Assessment of Real-Time ESP Failure Prediction Using Digital Twin, Machine Learning and Damage Modelling . *SPE Gulf Coast Section - Electric Submersible Pumps Symposium*.

Nadiah, Soim, S., & Sholihin. (2022). Implementation of Decision Tree Algorithm Machine Learning in Detecting Covid-19 Virus Patients Using Public Datasets. *Indonesian Journal of Artificial Intelligence and Data Mining (IJAIDM)*, 37 – 43.

Nasteski, V. (2017). An overview of the supervised machine learning methods. *Horizons*, 51-62.

Nguyen, N., Subervie, Y.-M., & Chong, J. (2020). Real-Time Automated Event Detection Framework for Electrical Submersible Pumps. *the Abu Dhabi International Petroleum Exhibition & Conference.* Abu Dhabi.

Panbarasan, M., Sankar, S., Venkateshbabu, S., & Balasubramanian, A. (2022). Characterization and performance enhancement of electrical submersible pump (ESP) using artificial intelligence (AI). *Materials Today: Proceedings*, 6864-6872.

Salem, A. M., Yakoot, M., & Mahmoud, O. (2022). A novel machine learning model for autonomous analysis and diagnosis of well integrity failures in artificial-lift production systems. *Advances in Geo-Energy Research*, 123-142.

Santoso, H. B., Vidrianto, M., Rahmawati, S. D., & Siagian, U. W. (2020). Predicting Failure in Electric Submersible Pump by Utilizing Machine Learning Based on Real-Time Sensor Data. *Proceeding Simposium IATMI 2020*.

Sharma, A., Songchitruksa, P., & Sinha, R. R. (2022). Integrating Domain Knowledge with Machine Learning to Optimize Electrical Submersible Pump Performance . *SPE Canadian Energy Technology Conference*.

Silvia, S., & Furlong, T. (2023). Case Study: Predicting Electrical Submersible Pump Failures Using Artificial Intelligence and Physics-Based Hybrid Models . *SPE Gulf Coast Section - Electric Submersible Pumps Symposium*.

Singh, B. K., Verma, K., & Thoke, A. S. (2015). Investigations on impact of feature normalization techniques on classifier's performance in breast tumor classification. *International Journal of Computer Applications*, 116.

Suyal, M., & Goyal, P. (2022). A Review on Analysis of K-Nearest Neighbor Classification Machine Learning Algorithms based on Supervised Learning. *International Journal of Engineering Trends and Technology*.

Tang, X., Wu, Z., Liu, W., Tian, J., & Liu, L. (2023). Exploring effective ways to increase reliable positive samples for machine learning-based urban waterlogging susceptibility assessments. *Journal of Environmental Management*.

Xiao, J. J., Shepler, R., Windiarto, Y., Parkinson, S., Fox, R., & Matlack, B. (2018). Development and Field Test of an Electric-Submersible-Pump Reliable-Power-Delivery System. *SPE Prod & Oper 33*, 449–458.

Zhen, X., Ning, Y., Du, W., Huang, Y., & Vinnem, J. E. (2023). An interpretable and augmented machine-learning approach for causation analysis of major accident risk indicators in the offshore petroleum industry . *Process Safety and Environmental Protection*, 922-933.