JOURNAL OF GAS TECHNOLOGY

Volume 10 / Issue 2 / Winter 2025 / Pages 43-62

Journal Home page: https://jgt.irangi.org

Analysis of Failures in Electrical Submersible Pumps Using Machine Learning Techniques

Mahya Faraji[1], Yasin Khalili[2], Mohammad Ahmadi[3*]

1. M.Sc., Department of Chemical Engineering, Institute of Petroleum Engineering, University of Tehran, Tehran, Iran
2. Ph.D. Student, Department of Petroleum and Geoenergy Engineering, Amirkabir University of Technology, Tehran, Iran
3. Associate Professor, Department of Petroleum and Geoenergy Engineering, Amirkabir University of Technology, Tehran, Iran

ARTICLE INFO

ORIGINAL RESEARCH ARTICLE

Article History:
Received: 22 September 2025
Revised: 04 October 2025
Accepted: 22 October 2025

Keywords:
Electrical Submersible Pumps (ESP)
Machine Learning
Predictive Maintenance
Anomaly Detection
Failure Prediction

DOR: 20.1001.1/jgt.2026.2078502.1064

ABSTRACT

Electrical Submersible Pumps (ESPs) play a critical role in enhancing production rates and operational efficiency in oil wells. However, their complex operating environments make them highly vulnerable to mechanical and electrical failures. Early detection of abnormal conditions and accurate prediction of pump performance are therefore essential for reducing downtime, minimizing operational costs, and extending pump lifespan.

This study investigates the application of machine learning techniques for identifying unstable operating conditions, detecting anomalies, and predicting failures in ESP systems. Using simulated monitoring data including motor temperature, intake temperature, current, and vibration several preprocessing and analytical methods were employed. Boxplot analysis was used to identify outliers, while Z-score analysis provided statistical detection of abnormal data points. K-means clustering and Principal Component Analysis (PCA) were implemented to reduce dimensionality, enhance visualization, and classify pump operating states into stable, unstable, anomalous, and failure modes.

Furthermore, a decision tree classifier was trained to determine the relative contribution of each parameter to pump failure. Model performance was evaluated using confusion matrices before and after applying PCA and clustering. Results show that incorporating these preprocessing techniques significantly improved classification accuracy from 0.9835 to 0.9967 for stable conditions and from 0.700 to 0.800 for failure conditions. The findings demonstrate the strong potential of machine learning methods for predictive maintenance of ESP systems and emphasize their value as a cost-effective strategy for enhancing reliability in the oil and gas industry.

How to cite this article
M. Faraji, Y. Khalili, M. Ahmadi. Analysis of Failures in Electrical Submersible Pumps Using Machine Learning Techniques. Journal of Gas Technology. 2025; 10(2): 43-62. (https://jgt.irangi.org/article_737830.html)

* Corresponding author.
E-mail address: m.ahmady@aut.ac.ir, (M. Ahmadi).

Available online 30 December 2025

1. Introduction

Electrical Submersible Pumps (ESPs) are among the most widely deployed artificial lift systems in the global oil and gas industry due to their ability to handle high production rates, operate in diverse reservoir conditions, and adapt to complex downhole environments. It is estimated that ESPs contribute to nearly 60% of global oil well production, underscoring their strategic importance in modern petroleum operations (Valbuena, Pereyra et al. 2016). Despite their widespread application, ESP systems are inherently vulnerable to premature failures caused by harsh downhole conditions, fluctuating flow regimes, and complex electro-mechanical interactions. Such failures often result in substantial production losses, expensive workover operations, and reduced overall field efficiency (Ladopoulos 2015). Consequently, improving ESP reliability and extending pump run life remain critical challenges for operators worldwide.

1.1. Importance and Economic Impact of ESP Failures

ESP failures represent a major source of non-productive time in oilfield operations. A single ESP failure can lead to prolonged production shutdowns, costly rig mobilization, and equipment replacement, with financial losses that may reach hundreds of thousands of dollars per event. In mature fields and deepwater applications, where intervention costs are particularly high, unexpected ESP failures can severely impact project economics. Furthermore, frequent failures reduce confidence in artificial lift system performance and complicate long-term production planning. These economic and operational consequences highlight the need for effective monitoring, early fault detection, and predictive maintenance strategies capable of identifying degradation before catastrophic failure occurs (Cardona, Vivas Sanchez et al. 2023).

1.2. Traditional ESP Failure Diagnosis Methods

Historically, ESP failure diagnosis has relied on physics-based models, threshold-based monitoring, and expert-driven rule systems derived from operational experience. Conventional approaches typically evaluate individual parameters such as motor temperature, motor current, vibration, or intake pressure against predefined operational limits (Abdelaziz, Lastra et al. 2017). While these methods are effective in detecting severe or already-developed faults, they exhibit several limitations. In particular, traditional techniques often struggle to identify early-stage degradation, capture nonlinear relationships among multiple operating variables, or adapt to changing reservoir and operating conditions (Peng, Feng et al. 2025). Moreover, such approaches require extensive expert knowledge and manual interpretation, limiting their scalability and consistency across different fields and well environments. These shortcomings have motivated the search for more adaptive and data-driven diagnostic solutions.

1.3. Machine Learning-Based Predictive Maintenance for ESP Systems

In recent years, machine learning (ML) techniques have gained increasing attention as powerful tools for ESP condition monitoring and predictive maintenance. Data-driven approaches leverage high-frequency sensor data such as temperature, current, vibration, and pressure to identify complex patterns associated with abnormal behavior and impending failure (Silvia, Gilad et al. 2023). Previous studies have demonstrated the potential of various ML algorithms, including artificial neural networks, support vector machines, decision trees, and ensemble models, to characterize ESP performance and predict failure events (Shi, Chen et al. 2019).

Despite these advances, many existing ML-based studies focus on single algorithms, binary classification between healthy and failed states, or limited subsets of operating conditions. In addition, issues such as systematic data preprocessing, handling of class imbalance, multi-state operational classification, and model interpretability are often insufficiently addressed. As a result, the practical deployment of ML models in real-world ESP monitoring systems remains challenging (Liang and El-Kadri 2018, Joseph, Adeoti et al. 2021, Li 2023).

1.4. Research Gap and Contribution

Despite the increasing application of machine learning techniques for ESP performance monitoring and failure diagnosis, several important limitations persist in the existing literature. Many prior studies rely on isolated ML models without integrated preprocessing frameworks, focus primarily on binary classification of normal versus failure conditions, or provide limited interpretability of model predictions. Furthermore, the combined use of statistical anomaly detection, unsupervised learning, dimensionality reduction, and explainable supervised classification has not been sufficiently explored within a unified framework for ESP condition assessment (Khalili, Rafiei et al. 2022, Ho, Fazilah et al. 2023, Kindi, Ghadani et al. 2024, Khalili, Ahmadi et al. 2025).

To address these gaps, this study proposes a hybrid and interpretable machine learning framework for ESP monitoring and failure prediction. The proposed approach integrates boxplot and Z-score-based statistical analysis for anomaly detection, K-means clustering for unsupervised identification of operational states, Principal Component Analysis (PCA) for dimensionality reduction and visualization, and decision tree classification for transparent and explainable failure prediction. Unlike previous studies, this framework enables multi-state classification of ESP behavior covering stable, unstable, anomalous, and failure conditions while demonstrating improved prediction accuracy through structured preprocessing. The proposed methodology provides a practical and extensible foundation for predictive maintenance of ESP systems and supports more reliable, data-driven operational decision-making.

2. Methodology

To identify unstable operating conditions, detect anomalies, and classify failure modes in ESPs, various machine learning and statistical analysis techniques were applied to the monitoring dataset (Lastra and Xiao 2022). These methods enable a structured assessment of operational parameters and support the development of predictive maintenance strategies (AlBallam 2022). This section introduces the analytical tools used in the study and their relevance to ESP performance evaluation.

2.1. Data Acquisition and Description

The dataset used in this study consists of simulated ESP operational monitoring data generated to reflect realistic field operating conditions and failure scenarios reported in the literature and manufacturer specifications. Simulation-based data were selected to enable controlled evaluation of the proposed methodology while ensuring consistency with documented ESP operating ranges and failure thresholds.

The dataset contains ESP pump data records, each corresponding to a distinct operational snapshot of the ESP system. Each record includes four key monitoring parameters: motor temperature, motor intake temperature, motor current, and motor vibration. Based on operating conditions, the data were labeled into four classes: stable operation, unstable operation, failure condition, and outlier (anomalous) state. The dataset exhibits class imbalance, with stable operation representing the majority of samples, reflecting realistic field behavior where failure events are relatively rare.

2.2. Data Preprocessing and Feature Engineering

Prior to model development, several preprocessing steps were applied to enhance data quality and robustness. Boxplot analysis was first employed to identify extreme outliers and assess data distribution characteristics. Subsequently, Z-score analysis was used for statistical anomaly detection, with data points exceeding $|Z| > 3$ flagged as abnormal.

All features were normalized to eliminate scale dependency and improve model convergence. Dimensionality reduction was performed using PCA to remove redundant correlations among variables and improve visualization and classification performance. No missing values were present in the simulated dataset; however, the preprocessing pipeline is designed to be extensible to field datasets where missing data handling may be required.

The ESP monitoring dataset be represented by a matrix

$$X = [x_{ij}] \in R^{N\times d} \quad (1)$$

where Nis the number of samples (records), dis the number of measured parameters (features), and x_{ij} denotes the value of feature j for sample i. In this study, d = 4, corresponding to motor temperature, intake motor temperature, motor current, and motor vibration.

Feature Normalization

To eliminate scale dependency across features, normalization was applied prior to clustering, PCA, and classification. Z-normalization (standardization) is defined as:

$$x_{ij}^{*} = \frac{x_{ij} - \mu_j}{\sigma_j} \quad (2)$$

where μj and σj are the mean and standard deviation of feature j, respectively.

2.3. Machine Learning Algorithms and Selection Rationale

A combination of unsupervised and supervised machine learning techniques was adopted to capture different aspects of ESP operational behavior. K-means clustering was selected for its simplicity and effectiveness in identifying latent operational states without prior labeling. PCA was applied to improve separability and reduce noise in the feature space.

For supervised classification, a decision tree classifier was employed due to its interpretability and suitability for engineering applications. Decision trees allow transparent identification of critical parameters influencing failure, which is essential for operational decision-making in ESP systems.

2.3.1. Machine Learning Algorithms and Selection Rationale

The boxplot is a widely used statistical visualization tool that provides an intuitive representation of data distribution. It illustrates key parameters such as the median, interquartile range (IQR), and extreme values (Bozhenko and Tatarnikova 2023, Krishnan, Kodamana et al. 2024). In the context of ESP monitoring, boxplot analysis serves two essential purposes:

1. Outlier Identification (Muslikh, Andono et al. 2023):

   Data points that lie significantly outside the normal operating range such as abrupt spikes in temperature or current may indicate sensor malfunction, transient anomalies, or early-stage pump degradation.

2. Initial Data Quality Assessment (Nandan Prasad 2024, Sobhy 2025):

   By visualizing the spread and skewness of each operational parameter, boxplots help identify inconsistencies, assess variability, and support subsequent preprocessing steps such as normalization and filtering.

This preliminary step ensures that the dataset fed into machine learning algorithms is representative, consistent, and free of extreme outliers that could bias model training.

In Box Plot, for each feature, the interquartile range (IQR) is computed as:

$$IQR = Q_3 - Q_1 \quad (3)$$

where Q_1 and Q_3 represent the first and third quartiles. A data point xis flagged as an outlier if it falls outside the lower and upper bounds:

$$x < Q_1 - 1.5\ IQR \text{ or } x > Q_3 + 1.5\ IQR. \quad (4)$$

2.3.2. Z-Score Method for Anomaly Detection

The Z-score, also known as the standard score, quantifies the number of standard deviations a data point deviates from the mean (Aggarwal, Gupta et al. 2019):

Importance in ESP Monitoring

A high absolute Z-score (e.g., |Z| > 3) indicates that the value is statistically anomalous. For ESP data:

1. High motor temperature or current may signal overloading or incipient failure
2. Sudden drops in intake temperature or pressure may indicate fluid entry issues
3. Large deviations in vibration readings may reflect mechanical imbalance or impending pump damage.

The Z-score method provides a straightforward and reliable approach for detecting anomalies before performing more advanced classification or clustering steps.

The Z-score (standard score) quantifies the deviation of a value from the mean in units of standard deviation:

$$Z = \frac{X - \mu}{\sigma} \quad (5)$$

where X is the observed value, μ is the mean, and σ is the standard deviation. A point is considered anomalous if:

$$|Z| > Z_{thr}, \quad (6)$$

where Z_{thr} is the anomaly threshold. In this study, $Z_{thr} = 3$was used.

2.3.3. K-Means Clustering

K-means is an unsupervised machine learning algorithm commonly used for partitioning datasets into groups (clusters) based on similarity. It iteratively assigns each data point to the nearest cluster center and updates the centroids until convergence (Oti, Olusola et al. 2021, Al Ghafri, Al Senani et al. 2025).

Relevance to ESP Analysis

In this study, K-means clustering was employed to categorize ESP operating states into:

1. Stable operation
2. Unstable operation
3. Faulty (failure-prone) conditions
4. Outlier or abnormal states

Grouping data into clusters allows the identification of operational patterns that are not explicitly labeled in the dataset. It is particularly useful in ESP monitoring where failure signatures may not be immediately apparent.

K-means partitions the dataset into Kclusters by minimizing the within-cluster sum of squares:

$$\min_{\{C_k\}_{k=1}^{K}} \sum_{k=1}^{K} \sum_{x_i \in C_k} \| \boldsymbol{x}_i - \boldsymbol{\mu}_k \|^2, \quad (7)$$

where C_k is the set of points assigned to cluster k, and μ_k is the centroid of cluster k:

$$\boldsymbol{\mu}_k = \frac{1}{|C_k|} \sum_{x_i \in C_k} \boldsymbol{x}_i. \quad (8)$$

Cluster assignment is performed by:

$$assign(\boldsymbol{x}_i) = \arg\min_{k \in \{1,\ldots,K\}} \| \boldsymbol{x}_i - \boldsymbol{\mu}_k \|^2. \quad (9)$$

2.3.4. Principal Component Analysis (PCA)

Principal Component Analysis (PCA) is a dimensionality reduction technique that transforms correlated variables into a smaller set of uncorrelated components (principal components). These components capture the maximum variance in the dataset (Peng, Han et al. 2020, Hasan, Abdulazeez et al. 2021, Song, Jun et al. 2024).

Benefits for ESP Performance Evaluation

1. Dimensionality Reduction:

   ESP datasets often include multiple correlated measurements. PCA simplifies the dataset while retaining essential information.

2. Visualization:

   By projecting data onto two principal components (PC1 and PC2), PCA makes it possible to visualize operational states and detect emerging patterns or clusters.

3. Noise Reduction:

   PCA helps remove redundant or noisy information, improving the performance of subsequent machine learning models.

   This technique is especially valuable in distinguishing between stable, unstable, and failure-prone operating regimes.

   PCA transforms the standardized feature matrix into a lower-dimensional representation that retains maximum variance. Let X^* denote the standardized data (Eq. 2). The covariance matrix is:

$$S = \frac{1}{N-1}(X^{*})^{\top} X^{*}. \quad (10)$$

PCA solves the eigenvalue problem:

$$Sv_m = \lambda_m v_m, \quad (11)$$

where λ_m and v_m are the eigenvalue and eigenvector corresponding to principal component m. The projection of a sample x_i^* onto the first pprincipal components is:

$$z_i = V_p^{\top} x_i^*, \quad (12)$$

where $V_p = [v_1, \ldots, v_p]$. The explained variance ratio for component mis:

$$EVR_m = \frac{\lambda_m}{\sum_{j=1}^{d} \lambda_j}. \quad (13)$$

2.3.5. Decision Tree Classification

Decision trees are supervised learning algorithms used to classify data based on a hierarchy of decision rules. Each internal node represents a test on a feature (e.g., motor temperature), and each leaf node corresponds to a predicted class (e.g., failure) (Castellanos, Serpa et al. 2020, Lasrado 2024).

Application to ESP Data

In this study, decision trees were used to:

1. Identify the relative importance of each parameter in predicting pump failure.

2. Create interpretable decision rules, such as "Motor temperature $> X \rightarrow$ High failure risk."

3. Classify new data into stable, unstable, or failure categories.

This method is particularly suitable for engineering applications because its outputs are transparent and easy to interpret, enabling engineers to understand the reasoning behind predictions.

A decision tree recursively splits the feature space to maximize class purity. Two commonly used impurity measures are Gini impurity and entropy.

For a node containing a class distribution $\{p_c\}_{c=1}^{C}$, Gini impurity is:

$$G = 1 - \sum_{c=1}^{C} p_c^2, \quad (14)$$

and entropy is:

$$H = -\sum_{c=1}^{C} p_c \log_2 (p_c). \quad (15)$$

For a candidate split that partitions the node into left and right child nodes, the impurity decrease (information gain) is computed as:

$$\Delta I = I(\text{parent}) - \left(\frac{n_L}{n} I(L) + \frac{n_R}{n} I(R)\right), \quad (16)$$

2.4. Model Training and Validation Framework

The dataset was divided into training and testing subsets using a 70/30 split. The models were trained on the training set and evaluated on unseen test data. Classification performance was assessed using confusion matrices, accuracy metrics, and class-wise evaluation to account for data imbalance.

Model performance was evaluated before and after applying PCA and K-means clustering to quantify the impact of preprocessing and dimensionality reduction. This framework ensures a fair and consistent comparison of model performance across different processing stages.

2.5. Performance Evaluation Metrics

To quantitatively evaluate the performance of the proposed machine learning framework, standard classification evaluation metrics were employed. Model predictions were assessed by comparing the predicted operating states with the true class labels using a confusion matrix. The confusion matrix provides a comprehensive summary of classification outcomes, including true positives, true negatives, false positives, and false negatives for each operating state (Al-Ballam, Karami et al. 2023; Khalili, Ahmadi et al. 2025).

From the confusion matrix, key performance indicators such as overall accuracy, class-wise accuracy, precision, and recall were derived. These metrics are particularly important in ESP failure prediction, where class imbalance is common and misclassification of failure states can lead to significant operational and economic consequences. Class-wise evaluation ensures that model performance is not dominated by the majority (stable operation) class and that failure and unstable conditions are adequately assessed.

In this study, confusion matrices were generated for the decision tree classifier under two different scenarios: (i) using the original feature space without dimensionality reduction and clustering, and (ii) after applying K-means clustering and PCA as preprocessing steps. This comparative evaluation framework allowed assessment of the impact of preprocessing and dimensionality reduction on classification performance, particularly in improving the identification of unstable and failure-prone operating states.

The confusion matrix compares predicted classes with true classes. For a given class, the following outcomes are defined: true positives (TP), false positives (FP), true negatives (TN), and false negatives (FN). Using these values, the main evaluation metrics are computed as:

Accuracy:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \quad (17)$$

Precision:

$$Precision = \frac{TP}{TP + FP} \quad (18)$$

Recall (Sensitivity):

$$Recall = \frac{TP}{TP + FN}. \quad (19)$$

F1-score:

$$F1 = 2 \cdot \frac{Precision \cdot Recall}{Precision + Recall}. \quad (20)$$

For multi-class classification, these metrics can be reported per class (class-wise) and averaged using macro- or weighted-averaging strategies to account for class imbalance.

3. Case Study: Reference ESP System

To evaluate the performance of the machine learning methods applied in this study, a representative ESP configuration was selected as a reference system. The case study is based on the operational specifications of a real industrial pump similar to the Baker Hughes FLEXPump™ Series 540, which is widely used in wells exhibiting diverse and challenging operating conditions.

The selected pump model is well-suited for medium-to-high productivity wells and offers enhanced resilience against temperature fluctuations, variable flow regimes, and harsh downhole environments. By comparing simulated failure scenarios with the technical specifications of this pump series, the study ensures that the analytical outcomes remain realistic and applicable to field operations.

3.1.Technical Specifications of the Reference Pump (FLEXPump™ Series 540)

The key operational characteristics of the selected pump are summarized as follows:

1. Pump Type

   A high-performance Electrical Submersible Pump designed for wells with complex downhole conditions, including variations in temperature, pressure, and inflow rate.

2. Installation Depth

   Capable of operating at depths up to 3000 meters, depending on reservoir pressure, fluid composition, and system configuration.

3. Production Capacity

   Designed to handle production rates in the range of 250 to 540 barrels per day (BPD), making it suitable for a broad spectrum of well productivity levels.

4. Operating Temperature Range

   The motor is engineered to tolerate temperatures up to 150°C, considerably higher than the simulated failure threshold of 110°C, providing sufficient thermal margin for safe and continuous operation.

5. Vibration Characteristics

   The pump design minimizes vibration, with operational vibration levels maintained below 1.5 RMS units, ensuring stable performance under varying load conditions.

6. Motor Power Rating

   Available in power configurations ranging from 5 to 50 horsepower, depending on well requirements and target production rates.

3.2. Advantages of the Selected Pump Configuration

The FLEXPump™ Series 540 exhibits several characteristics that make it suitable for predictive maintenance studies (Saxena , Werbinska-Wojciechowska and Rogowski 2025):

1. Enhanced Failure Resistance:

   The pump is specifically designed to withstand temperature and flow fluctuations that commonly lead to ESP failure.

2. Corrosion-Resistant Materials:

   Construction with anti-corrosive alloys improves reliability in wells containing H_2S, CO_2, or saline formation water..

3. Production Capacity

   The pump is compatible with advanced downhole monitoring systems capable of reporting real-time data on temperature, vibration, current, and intake pressure.

   These features make the pump an ideal reference system for evaluating the effectiveness of machine learning models in identifying failure conditions.

3.3. Alignment with Simulated Failure Conditions

The simulated Electrical Submersible Pump dataset used in this study was designed to represent realistic failure scenarios based on a combination of published literature, manufacturer technical specifications, and commonly reported ESP failure mechanisms. The operational thresholds defining failure conditions were selected to reflect values at which ESP systems typically experience performance degradation or imminent failure, as documented in prior studies and field experience.

Specifically, the failure-related operating conditions incorporated into the simulated dataset were defined as follows:

1. Motor temperature rising to approximately 110 °C, corresponding to elevated thermal stress levels commonly associated with insulation degradation and motor overheating in ESP systems, while remaining below the maximum rated motor temperature of the reference pump.
2. Motor current increasing to approximately 70 A, representing abnormal electrical loading conditions indicative of pump overload, flow restriction, or gas interference.
3. Vibration levels approaching 2.0 Root Mean Square (RMS) units, which exceed normal operating vibration ranges and are commonly linked to mechanical imbalance, impeller wear, or shaft misalignment.
4. Elevated intake motor temperature and non-uniform flow conditions, reflecting unstable inflow behavior, gas entry, or reservoir-induced flow disturbances that adversely affect pump performance.

These threshold values were intentionally selected to remain within the design limits of the reference Electrical Submersible Pump while representing conditions that are widely recognized as precursors to failure. As a result, the simulated anomalies are consistent with realistic field operating scenarios and allow meaningful comparison between simulated failure behavior and the technical tolerances of the reference pump system.

4. Results and Discussion

This section presents the results of the applied machine learning methods, including data preprocessing, anomaly detection, clustering, dimensionality reduction, feature importance analysis, and model evaluation (Figure 1). The findings demonstrate how each analytical step contributes to identifying operational patterns and predicting failures in ESP systems.

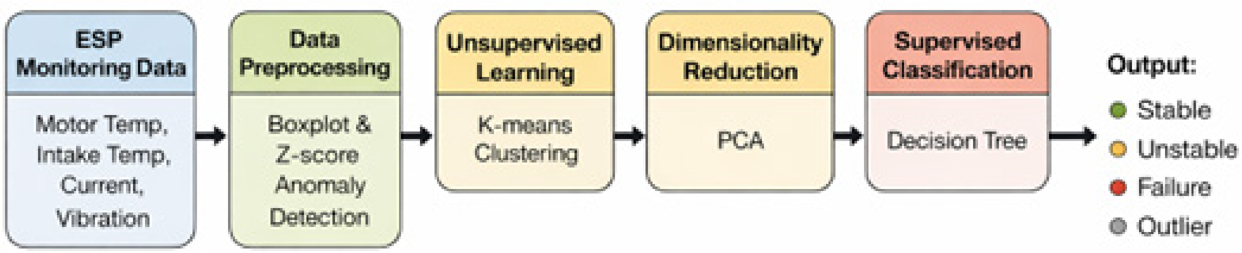

Figure 1. Overview of the Proposed Machine-Learning Framework for ESP Failure Detection and Classification

ESP monitoring data, including motor temperature, intake motor temperature, motor current, and vibration, are first preprocessed using boxplot analysis and Z-score-based anomaly detection. Unsupervised K-means clustering is then applied to identify latent operational states, followed by PCA for dimensionality reduction and visualization. Finally, a decision tree classifier is used to perform interpretable multi-state classification of ESP operating conditions into stable, unstable, failure, and outlier states.

4.1. Data Preprocessing Using Boxplot Analysis

Boxplot analysis was employed as an initial preprocessing step to visualize the distribution of the operational parameters and to identify statistical outliers. For each variable motor temperature, motor intake temperature, vibration, and motor current the boxplot reveals:

1. Minimum and maximum values
2. Median and interquartile range (IQR)
3. Potential outliers, which appear as isolated points beyond the whiskers

Key Observations

1. Detection of Outliers:

   The boxplots highlighted several extreme values, particularly in motor temperature and vibration. These outliers correspond to abnormal or failure-prone operating conditions and were considered during preprocessing.

2. Assessment of Data Variability:

   The spread of the quartiles indicated moderate-to-high variability in certain parameters especially vibration reflecting unstable wellbore flow or changing load conditions on the pump.

3. Improved Data Integrity:

   Identifying and isolating outliers prior to model training helps reduce noise and improves the robustness of downstream machine learning algorithms.

(Table 1) presents the statistical summary of the simulated ESP monitoring data obtained from the boxplot preprocessing step. These statistics provide an overview of the data distribution and support the identification of variability and potential outliers prior to further machine learning analysis.

Table 1. Statistical Summary Derived from Boxplot Preprocessing

Parameter	First Quartile (Q1)	Second Quartile (Median, Q2)	Third Quartile (Q3)	Maximum Value	Minimum Value	Mean
Motor Temperature (℃)	73.533	79.585	86.160	111.74	48.457	79.609
Motor Vibration (RMS)	1.005	1.468	1.845	3.0565	-0.10131	1.4846
Motor Current (A)	44.547	50.014	55.302	84.313	23.044	49.863
Motor Intake Temperature (℃)	67.754	74.917	81.287	106.24	39.615	74.608

(Figure 2) illustrates the distribution of key ESP operational parameters using boxplots, highlighting their central tendency, variability, and the presence of potential outliers.

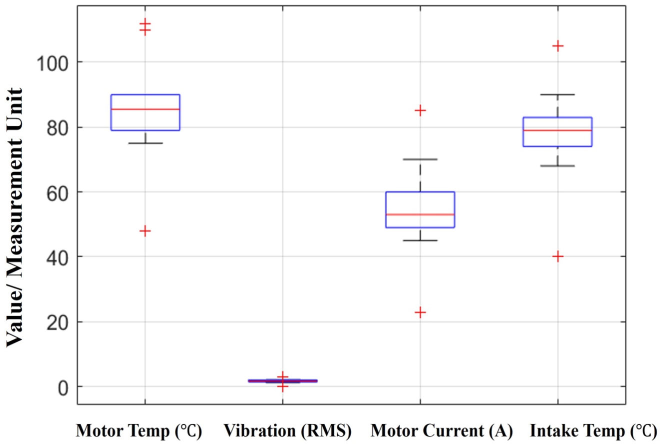

Figure 2. Boxplot of Simulated ESP Operational Parameters

Key Observations:

1. Boxplots revealed several extreme values notably in motor temperature and vibration.
2. Quartile spread indicated variability consistent with unstable or near-failure conditions.
3. Removing or flagging outliers enhanced the reliability of data used in later modeling.

4.2. Anomaly Detection Using Z-Score Analysis

Z-score analysis was implemented to detect statistically significant anomalies. Data points with $|Z| > 3$ were flagged as abnormal, which often corresponded to failure-related events.

The Z-score method effectively highlighted rare or extreme events across the monitoring period and provided a quantitative basis for classifying anomalies.

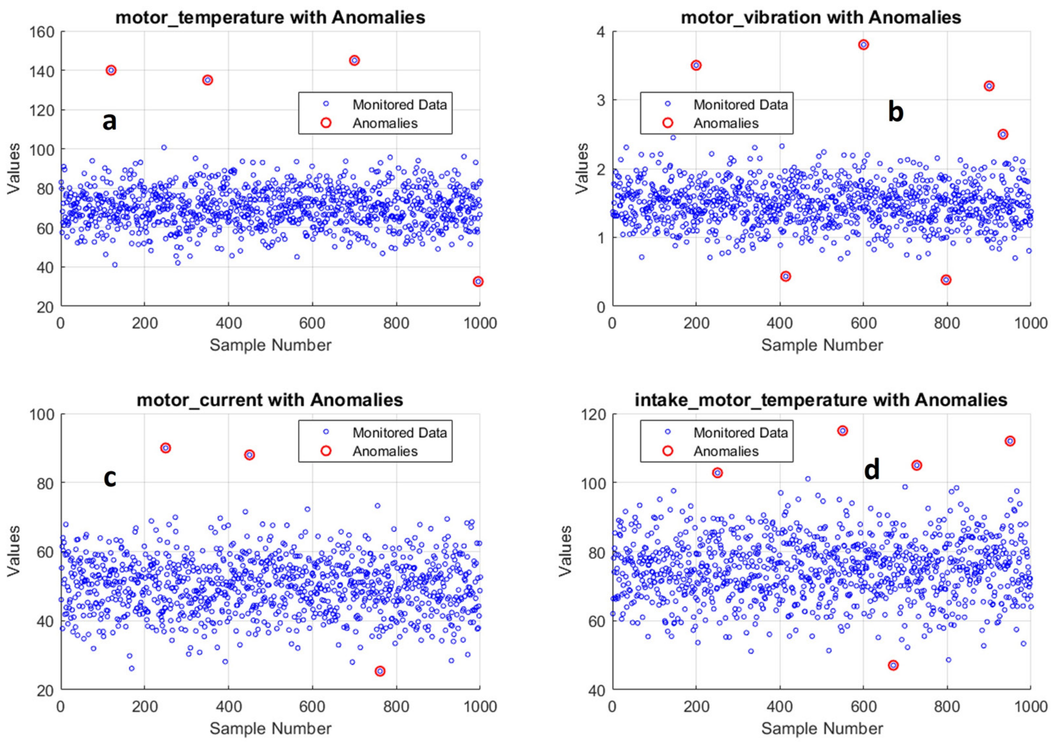

Figure 3. Z-score based Anomaly Detection of ESP Operational Parameters

(Figure 3) presents Z-score-based anomaly detection for ESP parameters, highlighting data points that deviate significantly from normal operating ranges.

(a) motor temperature, (b) motor vibration, (c) motor current, and (d) intake motor temperature. Blue markers represent monitored ESP operational data under normal conditions, while red markers indicate statistically significant anomalies identified using the Z-score method ($|Z| > 3$). The detected anomalies correspond to abnormal or failure-prone operating states associated with thermal overload, mechanical instability, electrical stress, or abnormal fluid inflow conditions.

As shown in Figure 2(a), motor temperature exhibits several high Z-score events corresponding to overheating conditions, which are commonly associated with motor overload and insulation degradation. Figure 2(b) highlights intake temperature anomalies that may reflect unstable inflow or fluid entry issues. Similarly, Figures 2(c) and 2(d) reveal abnormal current draw and elevated vibration levels, respectively, both of which are indicative of mechanical stress and incipient failure.

Findings:

1. Several high Z-score spikes in motor temperature indicated overheating patterns.
2. Elevated vibration anomalies were consistent with mechanical imbalance or impeller wear.
3. Sudden deviations in current corresponded to overload events.

4.3. Classification of Pump Operating States Using K-Means Clustering

K-means clustering was applied to group the data into operational regimes. The algorithm categorized the dataset into clusters corresponding to:

1. Stable operation
2. Unstable operation
3. Failure conditions
4. Outlier (abnormal) states

Cluster Interpretation

1. Stable Cluster:

   Represented by moderate temperatures (~80 °C), steady intake temperature (~75 °C), and low vibration levels.

   This group corresponds to normal, healthy ESP performance.

2. Unstable Cluster:

   Displayed elevated temperature and current values, indicating early deviation from optimal conditions.

3. Failure Cluster:

   Characterized by high motor temperatures (~110 °C), increased current, and abnormal vibration conditions typical of imminent pump failure.

4. Outlier Cluster:

   Includes extreme or unrealistic measurements, likely due to sensor errors or abrupt operational incidents.

   These clusters were subsequently used to enhance model interpretability and improve classification accuracy when combined with PCA.

(Figure 4) illustrates the clustering of ESP operational data into four distinct groups using the K-means algorithm, revealing patterns associated with stable, unstable, failure, and anomalous conditions.

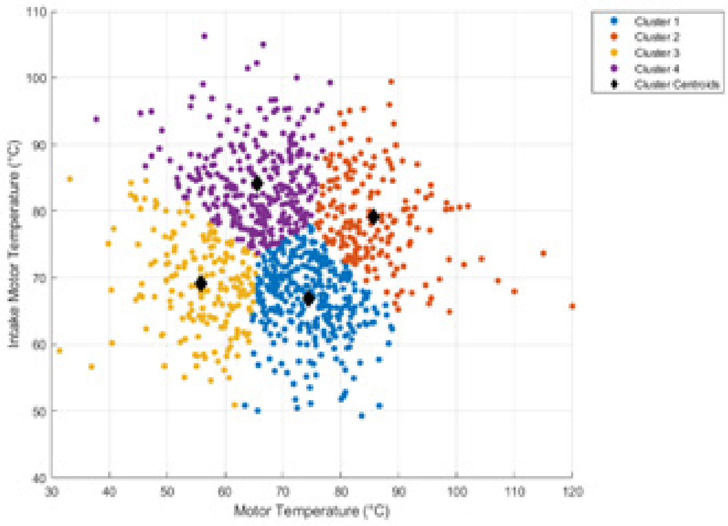

Figure 4. K-means Clustering of ESP Operational States

(Figure5) presents the time-series representation of ESP operational parameters, with K-means clustering used to distinguish stable, unstable, failure, and outlier operating conditions.

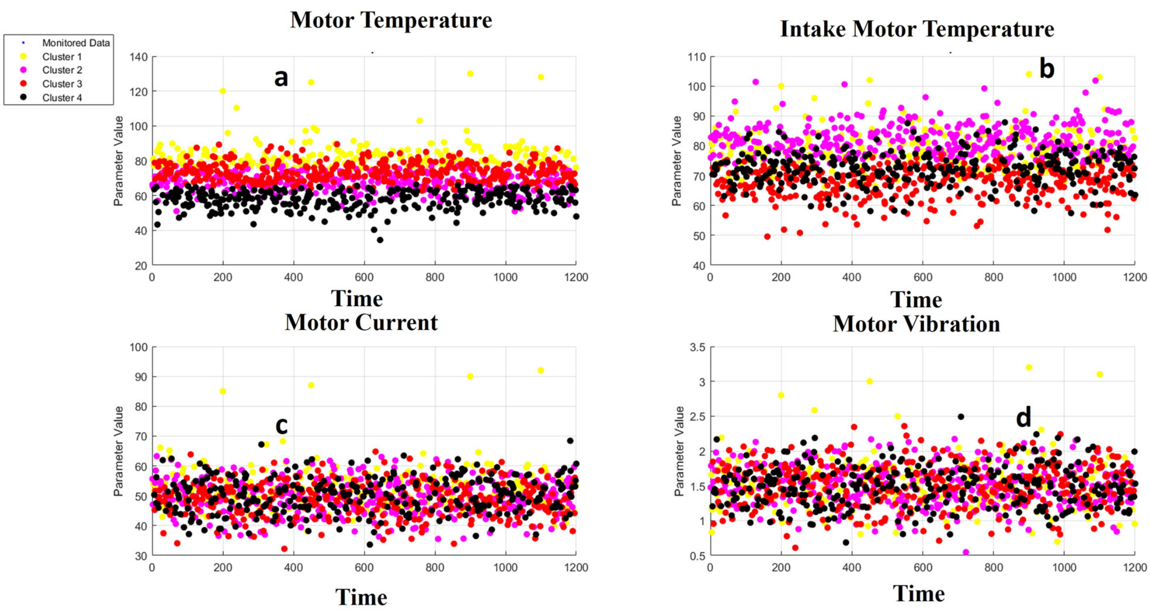

Figure 5. Time-series Representation of ESP Operational Parameters with K-means Cluster Labels

(a) Motor temperature,

(b) Motor intake temperature,

(c) Motor current,

(d) Motor vibration.

The color-coded clusters represent stable, unstable, failure, and outlier operating states identified through unsupervised learning.

This Figure illustrates the effectiveness of K-means clustering in separating ESP operational states. Stable operating conditions are characterized by moderate temperatures and low vibration levels, while failure-prone states exhibit elevated temperature, current, and vibration signatures. These visualizations support the subsequent classification results and demonstrate the ability of unsupervised learning to identify meaningful operational patterns.

Interpretation:

1. Stable cluster: moderate temperature and low vibration.
2. Unstable cluster: elevated temperature and current.
3. Failure cluster: high temperature (~110 °C), increased current, high vibration.
4. Outlier cluster: extreme sensor readings or rare transient events.

4.4. Dimensionality Reduction Using Principal Component Analysis (PCA)

Principal Component Analysis was applied to reduce the dataset to two principal components (PC1 and PC2), which together captured the majority of the variance.

Outcomes

1. Enhanced Visualization:

PCA allowed the projection of high-dimensional ESP data into a 2D space where clusters were clearly distinguishable.

2. Improved Cluster Separation:

Stable, unstable, failure, and outlier states formed distinct regions in the PCA-transformed space.

3. Noise Reduction:

PCA removed redundant correlations between variables, improving the performance of subsequent classifiers.

This step provided valuable insights into the relationships between operational parameters and facilitated clearer interpretation of pump condition states.

(Figure 6) illustrates the distribution of ESP operational states in PCA space, showing the separation between stable, unstable, failure, and outlier conditions.

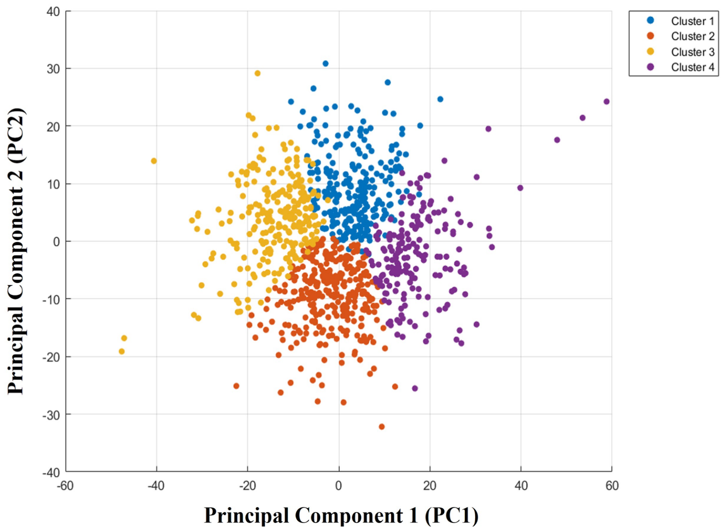

Figure 6. PCA-Based Visualization of ESP Operational States

Outcomes:

1. PCA clearly separated stable, unstable, failure, and outlier clusters.
2. Redundant correlations were removed, improving classifier performance.
3. PCA enhanced visual interpretability and noise reduction.

4.5. Feature Importance Analysis Using Decision Tree Classification

A decision tree classifier was trained using 70% of the data for training and 30% for testing. The model evaluated the relative contribution of each operational parameter to the prediction of failure.

Key Results

1. Motor temperature and motor intake temperature were identified as the most influential predictors of pump failure.
2. Motor current contributed moderately to decision splits.
3. Vibration had a lower but non-negligible influence, consistent with mechanical wear patterns.

These results align with real-world ESP failure mechanisms, where overheating and fluid entry instability are primary drivers of malfunction.

(Figure 7) shows the relative contribution of each ESP parameter to failure prediction, with motor temperature and intake temperature exhibiting the highest importance in the decision tree model.

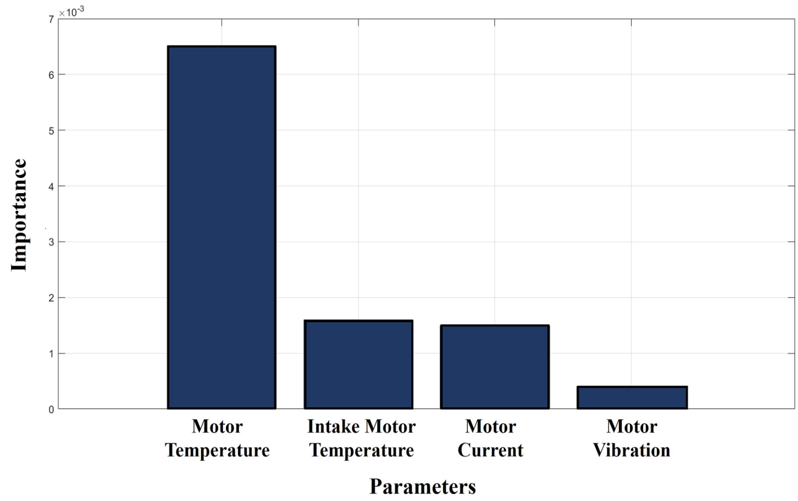

Figure 7. Feature Importance for Failure Classification

Key Results

1. Motor temperature and intake temperature contributed the highest importance.
2. Motor current played a secondary role.
3. Vibration influenced classification but to a lesser extent.

4.6. Model Evaluation Using Confusion Matrices

Confusion matrices were generated to evaluate classification performance before and after applying PCA and K-means-based preprocessing.

Before PCA and Clusterings

1. High accuracy in predicting stable states.
2. Reduced accuracy for unstable and failure states due to overlapping parameter ranges.

After PCA and Clustering

1. Stable condition accuracy improved from 0.9835 to 0.9967
2. Failure detection improved from 0.700 to 0.800
3. Slight improvement in unstable condition accuracy (0.9091 → 0.9095)

Interpretation

Applying PCA and clustering significantly enhanced the model's ability to classify failure states, demonstrating the importance of preprocessing and dimensionality reduction in predictive maintenance applications.

(Figure 8) presents the confusion matrix before applying PCA and K-means clustering, showing the model's classification performance across failure, outlier, stable, and unstable operating states.

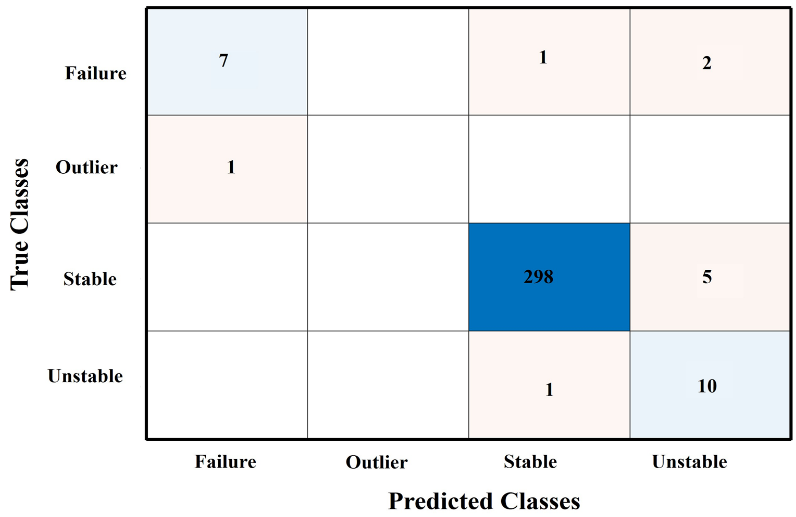

Figure 8. Confusion Matrix Before PCA and K-means Clustering

(Figure 9) presents the confusion matrix after applying PCA and K-means clustering, showing improved classification accuracy, particularly for stable and failure operating states.

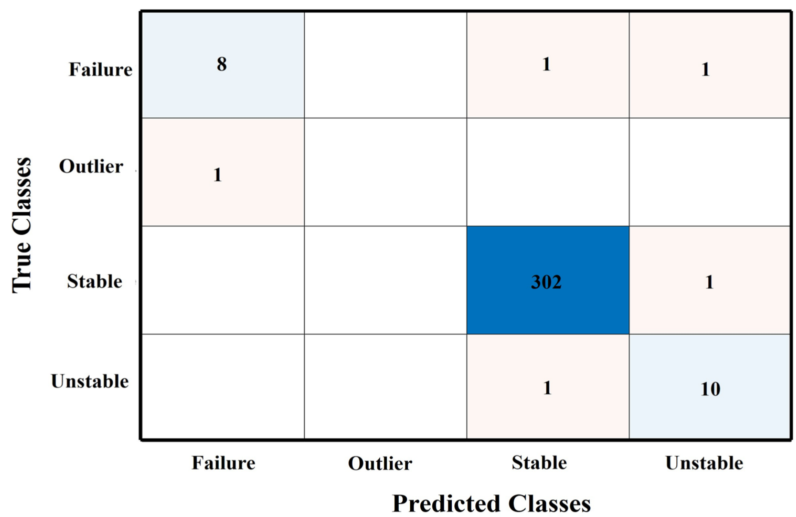

Figure 9. Confusion Matrix After PCA and K-means Clustering

(Table 2) summarizes the classification accuracy for stable, unstable, and failure states before and after preprocessing, demonstrating the improvement achieved through PCA and K-means clustering.

Table 2. Accuracy Values Before and After Preprocessing

Predicted Class	Accuracy Before Preprocessing	Accuracy After Preprocessing
Stable	0.9835 (98.35%)	0.9967 (99.67%)
Unstable	0.9091 (90.91%)	0.9095 (90.95%)
Failure	0.7000 (70.00%)	0.8000 (80.00%)

5. Future Outlook

While the findings of this study are promising, several opportunities exist for expanding and refining the proposed framework:

1. Use of Real Field Data

   Future work will focus on applying the proposed framework to field-operational ESP datasets obtained from producing wells. Validation using real-world data will enable assessment of model robustness under varying reservoir conditions, sensor noise, and long-term operational variability, further enhancing the practical applicability of the proposed approach.

2. Integration of Advanced Algorithms

   More sophisticated approaches such as Random Forests, Gradient Boosting, Support Vector Machines, or Deep Learning may further improve anomaly detection and failure prediction accuracy.

3. Real-Time Monitoring and Deployment

   Implementing the trained models in a real-time monitoring environment would enable continuous health assessment and automated alert systems for proactive well management..

   (Table 1) presents the statistical summary of the simulated ESP monitoring data obtained from the boxplot preprocessing step. These statistics provide an overview of the data distribution and support the identification of variability and potential outliers prior to further machine learning analysis.

4. Multivariate Time-Series Analysis

   Incorporating temporal dependencies through methods such as LSTM networks or ARIMA models could capture dynamic ESP behavior and improve early failure detection.

5. Digital Twin Development

   Incorporating temporal dependencies through methods such as LSTM networks or ARIMA models could capture dynamic ESP behavior and improve early failure detection.

6. Conclusion

This study presented a structured and interpretable machine-learning framework for identifying unstable operating conditions, detecting anomalies, and predicting failure states in Electrical Submersible Pump (ESP) systems. The framework integrates statistical preprocessing, unsupervised learning, dimensionality reduction, and supervised classification, and was evaluated using a simulated ESP monitoring dataset designed to reflect realistic field operating conditions.

The results demonstrate that data preprocessing plays a critical role in improving model performance and reliability. Boxplot analysis and Z-score-based anomaly detection successfully isolated abnormal observations associated with thermal, electrical, and mechanical deviations, thereby improving the quality of data supplied to subsequent learning stages.

Unsupervised K-means clustering enabled effective segmentation of ESP operational behavior into stable, unstable, failure, and outlier states without prior labeling. This clustering step provided a clear operational structure that enhanced the interpretability of the dataset and supported multi-state condition assessment, which is more representative of real ESP behavior than binary classification.

Principal Component Analysis (PCA) reduced the dimensionality of the dataset while preserving the dominant variance, leading to improved separation of operating regimes in the reduced feature space. The application of PCA contributed directly to improved classification performance and clearer visualization of degradation patterns.

The decision tree classifier identified motor temperature and intake motor temperature as the most influential parameters for ESP failure prediction. This finding is consistent with established ESP failure mechanisms related to overheating, insufficient cooling, and unstable inflow conditions, confirming the physical relevance of the machine-learning results.

Quantitatively, the integration of clustering and PCA resulted in measurable improvements in classification accuracy. The accuracy of stable-state identification increased from 0.9835 to 0.9967, the unstable-state accuracy improved to 0.9095, and the failure-state accuracy increased from 0.700 to 0.800. These results demonstrate that structured preprocessing and multi-stage learning significantly enhance the detection of failure-prone conditions, particularly for minority classes.

Overall, this study confirms that machine learning provides a practical, accurate, and interpretable approach for ESP predictive maintenance. The proposed framework enables earlier identification of abnormal operating conditions, supports informed operational decision-making, and has the potential to reduce unplanned workovers, extend ESP run life, and improve field-level productivity. While the present work is based on simulated data, the methodology is directly transferable to field-operational datasets and provides a solid foundation for future real-time ESP monitoring and failure prediction applications.

References

Abdelaziz, M., R. Lastra and J. Xiao (2017). ESP data analytics: Predicting failures for improved production performance. Abu Dhabi International Petroleum Exhibition and Conference, SPE.

Aggarwal, V., V. Gupta, P. Singh, K. Sharma and N. Sharma (2019). Detection of spatial outlier by using improved Z-score test. 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), IEEE.

Al Ghafri, A. H., A. Al Senani, H. Al-Salmi, S. Al-Mamari and E. Artun (2025). An Unsupervised Learning Framework for Managing ESP-Driven Oil Field Operations with Water Injection in Northern Oman. SPE EOR Conference at Oil and Gas West Asia, SPE.

AlBallam, S. (2022). "Applying machine learning models to diagnose failures in electrical submersible pumps."

Bozhenko, V. and T. Tatarnikova (2023). Application of data preprocessing in medical research. 2023 Wave Electronics and its Application in Information and Telecommunication Systems (WECONF), IEEE.

Cardona, L., P. Vivas Sanchez and B. Joya (2023). Failure prediction methodology for ESP and operational behavior. SPE Latin America and Caribbean Petroleum Engineering

Conference, SPE.

Castellanos, M. B., A. L. Serpa, J. L. Biazussi, W. M. Verde, N. d. S. D. A. J. J. o. P. S. Sassim and Engineering (2020). "Fault identification using a chain of decision trees in an electrical submersible pump operating in a liquid-gas flow." 184: 106490.

Hasan, B. M. S., A. M. J. J. o. S. C. Abdulazeez and D. Mining (2021). "A review of principal component analysis algorithm for dimensionality reduction." 2(1): 20-30.

Ho, Y. H., S. R. Fazilah, J. Nava Bastidas, A. A. Amsidom, E. A. Rosland, K. N. Idris, R. Masoudi, K. Trjangganung, Z. Mohd Khair and M. S. majid (2023). Shifting focus to electrical submersible pump (ESP) technology to enhance well productivity and recovery in Malaysia fields. International Petroleum Technology Conference, IPTC.

Joseph, A., A. J. J. o. A. S. Adeoti and E. Management (2021). "Impact of fluid properties on Electric Submersible Pumps (ESP) performance and run life in a well." 25(2): 139-143.

Khalili, Y., M. Ahmadi, M. K. J. A. J. f. S. Moraveji and Engineering (2025). "A comprehensive review of failure modes in electrical submersible pumps: Diagnosis, predictive maintenance, and engineer's guide." 1-22.

Khalili, Y., Y. Rafiei and M. Sharifi (2022). "Reservoir Characterization by Applying Pressure Transient Analysis on Data Obtained from Electrical Submersible Pumps."

Kindi, M., A. Ghadani, S. Shekaili, F. Aufi, S. Al Wahaibi and C. Moreno Gomez (2024). Revolutionizing Oil Lift Operations: Harnessing Data Analytics for Esp Energy Management, Cost Savings, Co_2 Emissions Reduction, and Enhanced Pump Runlife. International Petroleum Technology Conference, IPTC.

Krishnan, N. A., H. Kodamana and R. Bhattoo (2024). Data Visualization and Preprocessing. Machine Learning for Materials Discovery: Numerical Recipes and Practical Applications, Springer: 25-46.

Ladopoulos, E. J. U. J. H., 3 , 1 (2015). "Four-dimensional petroleum exploration & non-linear esp artificial lift by multiple pumps for petroleum well development." 14.

Lasrado, V. (2024). ESP Predictive Analytic Using a Single Application with Machine Learning and Fault Tree Models. SPE Middle East Artificial Lift Conference and Exhibition, SPE.

Lastra, R. A. and J. Xiao (2022). Machine learning engine for real-time ESP failure detection and diagnostics. SPE Middle East Artificial Lift Conference and Exhibition, SPE.

Li, L. J. E. (2023). "Energy Consumption Analysis and Optimization of Electric Submersible Pump System." 15(5): 269-274.

Liang, X. and A. El-Kadri (2018). Factors affecting electrical submersible pump systems operation. 2018 IEEE Electrical Power and Energy Conference (EPEC), IEEE.

Muslikh, A. R., P. N. Andono, A. Marjuni and H. A. Santoso (2023). Systematic literature review of data distribution in preprocessing stage with focus on outliers. 2023 International Seminar on Application for Technology of Information and Communication (iSemantic), IEEE.

Nandan Prasad, A. (2024). Data Quality and Preprocessing. Introduction to Data Governance for Machine Learning Systems: Fundamental Principles, Critical Practices, and Future Trends, Springer: 109-223.

Oti, E. U., M. O. Olusola, F. C. Eze and S. U. J. c. Enogwe (2021). "Comprehensive review of K-Means clustering algorithms." 12(08): 22-23.

Peng, L., L. Feng, Q. Guang, J. Zhu, H. Liu, Z. Nie, C. Ma, C. Di and Q. Wu (2025). Real-Time ESP Management Framework Using Hybrid Physics-Based and ML Models. SPE Offshore

Europe Conference and Exhibition, SPE.

Peng, L., G. Han, A. L. Pagou, J. J. J. o. P. S. Shu and Engineering (2020). "Electric submersible pump broken shaft fault diagnosis based on principal component analysis." 191: 107154.

Saxena, A. "Application of Data Science in Pump Maintenance."

Shi, J., S. Chen, X. Zhang, R. Zhao, Z. Liu, M. Liu, N. Zhang and D. Sun (2019). Artificial lift methods optimising and selecting based on big data analysis technology. International Petroleum Technology Conference, IPTC.

Silvia, S., Y. Gilad, T. A. Wilson, B. Akbari and E. R. Furlong (2023). Case study: predicting electrical submersible pump failures using artificial intelligence and physics-based hybrid models. SPE Middle East Intelligent Oil and Gas Symposium, SPE.

Sobhy, M. A. (2025). "Detecting Electrical Submersible Pump (ESP) Failures and Estimating Run Life Using Artificial Neural Networks."

Song, Y., S. Jun, T. C. Nguyen, J. J. J. o. P. E. Wang and P. Technology (2024). "Experimental data-driven model development for ESP failure diagnosis based on the principal component analysis." 14(6): 1521-1537.

Valbuena, J., E. Pereyra and C. Sarica (2016). Defining the artificial lift system selection guidelines for horizontal wells. SPE Artificial Lift Conference and Exhibition-Americas, SPE.

Werbinska-Wojciechowska, S. and R. J. S. Rogowski (2025). "Proactive Maintenance of Pump Systems Operating in the Mining Industry—A Systematic Review." 25(8): 2365.

تحلیل خرابی پمپ‌های شناور الکتریکی با استفاده از تکنیک‌های یادگیری ماشین

• محیا فرجی[۱]، یاسین خلیلی[۲]، محمد احمدی[۳*]

۱. دانشجو کارشناسی ارشد، دانشکده مهندسی شیمی، انستیتو مهندسی نفت، دانشگاه تهران، ایران

۲. دانشجوی دکترا، دانشکده مهندسی نفت و زمین انرژی، دانشگاه صنعتی امیرکبیر، تهران، ایران

۳. دانشیار، دانشکده مهندسی نفت و زمین انرژی، دانشگاه صنعتی امیرکبیر، تهران، ایران

(ایمیل نویسنده مسئول: m.ahmady@aut.ac.ir)

چکیـــــده

پمپ‌هـای شـناور الکتریـکی نقـش بسیـار مهـمی در افزایـش نـرخ تولیـد و بهبـود کارایی عملیـاتی چاه‌های نفت ایفـا می‌کننـد. با این حـال، شـرایط پیچیـده و چالش‌برانگیـز محیـط درون‌چـاهی باعـث می‌شـود ایـن پمپ‌هـا در معـرض خرابی‌هـای مکانیـکی و الکتریـکی قـرار گیرنـد. ازایـن‌رو، شناسـایی زودهنـگام وضعیت‌هـای غیرعـادی و پیش‌بیـنی دقیـق عملکـرد پمـپ بـرای کاهـش زمـان توقـف تولید، کمینه‌سـازی هزینه‌هـای عملیـاتی و افزایـش طـول عمـر تجهیـز ضـروری اسـت.

در ایـن پژوهـش، کاربـرد روش‌هـای یادگیـری ماشین برای شناسـایی وضعیت‌هـای ناپایدار، تشـخیص ناهنجاری‌هـا و پیش‌بینی خرابی در سیسـتم‌های پمـپ شـناور الکتریـکی مـورد بررسی قرار گرفته اسـت. با اسـتفاده از داده‌های شبیه‌سازی‌شـده شـامل دمـای موتور، دمای ورودی موتـور، جریـان و لـرزش موتـور، چنديـن روش پیش‌پردازش و تحلیل داده به‌کار گرفته شـد. تحلیل نمودار جعبه‌ای جهت شناسـایی داده‌هـای پـرت اسـتفاده شـد و روش زد-اسـکور نیز تشـخیص آمـاری نقاط غیرعـادی را امکان‌پذیـر کرد. همچنیـن الگوریتم خوشـه‌بندی کا-میانگیـن و تحلیـل مؤلفه‌هـای اسـاسی بـرای کاهـش ابعـاد داده، بهبـود تجسـم و دسـته‌بندی وضعیت‌هـای پمپ بـه حالت‌هـای پایدار، ناپایـدار، ناهنجـار و خرابی مورد اسـتفاده قـرار گرفت.

علاوه بـر ایـن، یـک طبقه‌بنـدی درخـت تصمیم‌گیـری بـرای تعییـن میزان تأثیر نسـبی هـر پارامتـر در بروز خـرابی آموزش داده شـد. عملکـرد مـدل بـا اسـتفاده از ماتریس سـردرگمی، قبـل و بعد از اعمـال روش‌های آنالیز مؤلفه اسـاسی و خوشـه‌بندی ارزیـابی گردید. نتایج نشـان می‌دهـد کـه اعمـال ایـن روش‌هـا دقـت طبقه‌بنـدی را به‌طـور قابل‌توجـهی بهبـود داده اسـت؛ به‌طـوری که دقـت شناسـایی حالت پایـدار از ۰/۹۸۳۵ بـه ۰/۹۹۶۷ و دقـت تشـخیص حالـت خـرابی از ۰/۷ بـه ۰/۸ افزایـش یافتـه اسـت. یافته‌هـای ایـن پژوهش نشـان‌دهنده پتانسیـل بـالای روش‌هـای یادگیـری ماشیـن در نگهداشـت پیش‌بینانه پمپ‌های شـناور الکتریکی بوده و بـر ارزش آن‌ها به‌عنـوان راهبردی مقرون‌به‌صرفـه بـرای افزایـش قابلیـت اطمینـان در صنعـت نفـت و گاز تأکید دارد.

واژگان کلیدی: پمپ‌های شناور الکتریکی، یادگیری ماشین، مراقبت پیش‌بینانه، تشخیص ناهنجاری، پیش‌بینی خرابی