UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

APPLYING MACHINE LEARNING MODELS TO DIAGNOSE FAILURES IN ELECTRICAL SUBMERSIBLE PUMPS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

SHAIKHAH SHAKER ALBALLAM

Norman, Oklahoma

2022

APPLYING MACHINE LEARNING MODELS TO DIAGNOSE FAILURES IN ELECTRICAL SUBMERSIBLE PUMPS

A DISSERTATION APPROVED FOR THE

MEWBOURNE SCHOOL OF PETROLEUM AND GEOLOGICAL ENGINEERING

BY THE COMMITTEE CONSISTING OF

Dr. Hamidreza Karami (Chair)

Dr. Deepak Devegowda (Co-Chair)

Dr. Charles D. Nicholson

Dr. Eduardo Pereyra

Dr. Xingru Wu

Dedication

This dissertation is dedicated to my mother, Lamiaa AlSheetan, for her endless love, support, and encouragement.

Acknowledgments

In the name of Allah, the Most Gracious, the Most Merciful

First, I would like to thank Almighty ALLAH for his support from the beginning till the time of publishing the dissertation.

Second, I would like to express my personal appreciation to the following people and organizations for their support and assistance during my Ph.D. studies:

Kuwait Petroleum Corporation (KPC): For granting me the Scholarship and the opportunity to pursue my Ph.D. studies.

Kuwait Oil Company (KOC): For supporting my research with field data.

Prof. Eissa Al-Safran (Kuwait University): For his usual support and guidance.

Mewbourne School of Petroleum and Geological Engineering, University of Oklahoma: For their resources and academic support.

Dr. Hamidreza Karami: I thank him for his input, unceasing guidance, and the invaluable help of constructive comments and suggestions throughout the development of this dissertation.

Dr. Deepak Devegowda: I thank him for his assistance and invaluable advice during the preparation of my dissertation.

Finally, I appreciate all my committee members for the time and effort devoted to my dissertation. Their valuable comments on the dissertation made my work more focused.

Table of Contents

List of Tables

List of Figures

Abstract

Electrical Submersible Pump (ESP) failures are unanticipated but common occurrences in oil and gas wells. It is necessary to detect the onset of failures early and prevent excessive downtime. This study proposes a novel approach utilizing multi-class classification machine learning models to predict various ESP specific failure modes (SFM's). A comprehensive dataset and various machine learning algorithms are utilized. The prediction periods of 3 hours to 7 days before the failure are evaluated to minimize false alarms and predict the true events.

The ML models are based on field data gathered from surface and downhole ESP monitoring equipment over five years of production of 10 wells. The dataset includes the failure cause, duration of downtime, the corresponding high-frequency pump data, and well production data. According to these data, most ESP operational failures are characterized as electrical failures.

Four modeling designs are used to handle the data and transform them into actionable information to predict various ESP failure modes at different prediction periods. Several ML models are tested and evaluated using precision, recall, and F1-score performance measures. The K-Nearest Neighbor (KNN) model outperforms the other algorithms in forecasting ESP failures. Some other tested models are Random Forest (RF), Decision Tree (DT), Multilayer Perceptron (MLP) Neural Network, etc. The findings of these ML models reveal that as the prediction period extends beyond three days, it becomes more challenging to predict the true failures. Furthermore, all tested designs show similarly good performances in predicting ESP specific failures. The design

that integrates the impacts of gas presence and pump efficiency while minimizing the number of input variables is suggested for general use.

Based on the field data, a Weibull model is built to estimate the probability of failure. The mean time between failure (MTBF) values are utilized as inputs to the Weibull analysis. The Weibull shape and scale parameters are estimated using Median Rank Regression. Then the Weibull Probability plots are generated with high R^2 values (86.5-99.4%) and a low p-value for all wells. The results show increases in pump unreliability with time for all the wells.

By integrating the outcomes of the ESP Failure prediction ML model with the Weibull unreliability model, a powerful tool is provided. This tool allows the engineers to detect failures early, diagnose potential causes, and propose preventive actions. It is crucial in aiding the operators in transitioning from reactive to proactive and predictive maintenance of artificial lift operations.

Chapter 1 Introduction

Various Artificial Lift technologies (AL) are used to raise a hydrocarbon well's production rate and alter its bottomhole pressure. Nearly half of the world's 2 million oil wells are supported by artificial lift methods, demonstrating the widespread use of this technology (Lea, 2007). According to Mordor Intelligence (2022), recent market research indicates that the AL systems will expand at a compound annual growth rate (CAGR) of around 4.5% between 2022 and 2027. As of 2021, a significant portion of global oil and gas production comes from mature fields, necessitating an artificial lift system to improve the recovery rates (Takacs, 2017). Asia-Pacific is among the fastest-growing AL markets, while North America led the global AL systems market in 2021 (Mordor Intelligence, 2022). As shown in Figure 1-1, North America is likely to dominate the market due to its many mature oil and gas reserves and expanding energy demand.

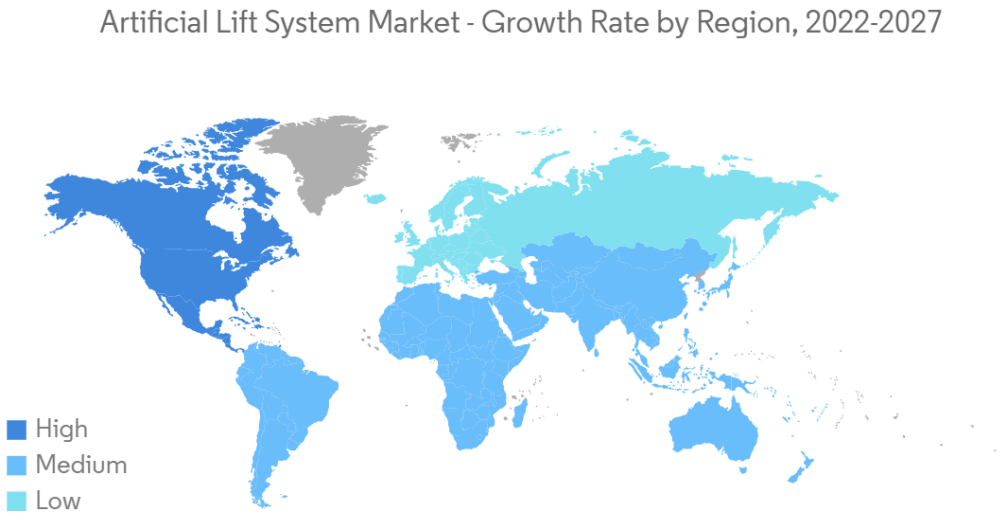

Figure 1-1: Artificial Lift Systems Predicted Growth Across the World (Mordor Intelligence, 2022)

Electrical submersible pumps (ESPs) are the best option for boosting production rates in major oil wells that have reached or exceeded their peak oil productions. ESPs are a common artificial lift method in the petroleum industry, particularly in high-flow rate oil wells. They are anticipated to dominate the market owing to their adaptability and technological advancements. However, ESP failures are unanticipated and common occurrences in the oilfield, with significant financial impacts for the operators. ESPs made up 49% of the $10.9 billion artificial lift industry in 2012, making their failures a significant financial loss owing to lost production and replacement or intervention costs (Donner *et al.*, 2014). GE Oil & Gas conducted a study, assessing the yearly cost of lost production and intervention due to ESP shutdowns at $3 million and $1 million per well per year, respectively (Carrillo, 2013). Developing analytical models to detect and mitigate ESP failures is critical in lowering the downtime and extending the life of an ESP.

This study uses an extensive dataset from several years of production of multiple wells to train a comprehensive set of machine learning algorithms and predict ESP failures before they happen. The results of the proposed data-driven models are crucial in aiding the operators toward transitioning from reactive event-based maintenance to proactive and predictive maintenance of artificial lift operations. This approach increases the life and efficiency of the ESP by boosting the uptime, lowering the intervention costs, and optimizing the production.

1.1 Objectives

This study has the following main objectives:

- Conduct a comprehensive literature review to determine the development of ESP failure prediction through time and the systems necessary to assist in the process of predicting failures at early stages.

- Conduct a statistical analysis on the field data gathered from the surface and downhole ESP monitoring equipment to classify and characterize the mechanisms of ESP failure.
- Construct machine learning models (ML) to handle the high frequency ESP data and predict various ESP failure modes at different prediction periods. Utilize performance metrics to validate the models.
- Insert the physical aspects of ESP operations in the developed machine learning model to provide a hybrid approach for failure prediction. Identify the best model with the highest accuracy in forecasting ESP failures.

1.2 Scope of Work

This study uses the field data acquired from Kuwait Oil Company (KOC) assets, including over five years of production from ten operational wells. The recorded data involves different field conditions, including dynamic, static, and historical data. The dynamic high-frequency (per minute) data from surface and downhole ESP monitoring equipment are collected (variable speed drive, pump, and wellhead). The static data include the well completion, ESP design, and reservoir fluid information. The data were collected from wells with high ESP failure rates to investigate the general ESP failures and their specific failure modes. Additionally, historical operational data are included in the study.

The first step in constructing a machine learning model is data preparation to clean, organize, and categorize raw data prior to use by the model. Then comes the pre-processing step, which involves managing missing values, detecting anomalies, splitting the data into training and testing sets, normalizing the data, and optimizing hyperparameters. These steps account for 80% of the overall time. Consequently, the tuned parameters are used to construct the new ML model.

This model will then be evaluated using the test dataset in order to assess the performance of the ML algorithms.

In addition, generating several features for the machine learning model is essential for developing a hybrid approach to failure prediction that accounts for the physical aspects of ESP operations. This study investigates the gas interference in the ESP system and evaluates pump performance based on field data. Additionally, a detailed technique is described for estimating the fluid properties based on the field data and using it to assess gas interference in the pump.

The specific failure mode (SFM) is predicted utilizing ESP-labeled datasets as inputs. A supervised multiclass classification model is used. Several ML models are developed to handle the high frequency of ESP operational data and predict the ESP failure modes throughout prediction periods. The best model with the highest accuracy in predicting ESP failures is then determined based on the performance measures.

1.3 Outline

This work is divided into eight chapters. To begin, Chapter 2 contains a thorough review of the literature on the evolution of ESP failure prediction through time, as well as a detailed review of ESP pump performance and gas interference. Chapter 3 describes the ESP failures statistical analysis and a detailed ESP failure overview utilizing field data. Then, the approach for quantifying the probability of failure, the economic effect, and the ensuing NPV losses are presented in Chapter 4. Chapter 5 contains a full discussion of the data analytics approaches implemented to diagnose and predict the SFMs. This chapter discusses the various designs constructed and the various applied machine learning algorithms. Chapter 6 summarizes and contrasts the results of the machine learning models for each developed design. Chapter 7 provides

an overview of the future implementations of machine learning models in the live field conditions. Finally, Chapter 8 concludes this study by highlighting the significant findings and recommends some noteworthy subsequent works that might potentially improve the research.

Chapter 2 Literature Review

This chapter includes a review of the available literature on this study's primary research topic. First, a general review is conducted on Electrical Submersible Pumps (ESPs) as one of the most widely used artificial lift techniques in the petroleum industry. Then, a thorough review is provided of the literature on the evolution of ESP failure prediction through time, followed by a detailed review of ESP pump performance and gas interference calculations.

2.1 Electrical Submersible Pump

With the ever-declining production rates from oil and gas fields worldwide, the demand for efficient artificial lift techniques is increasing every day to enhance the production (Igwilo *et al*., 2018). The selection of a suitable artificial lift technique for a given well is a function of various operational parameters (Nguyen, 2020). The goal is always to maximize the profit from the well by increasing the production while maintaining the artificial lift-related expenses and downtime to a minimum. This is what makes a thorough understanding of an artificial lift technique's operating ranges, limitations, and failures vital prior to its application.

Electrical Submersible Pumps (ESPs) are one of the most widely used artificial lift techniques in the petroleum industry, especially for highly productive oil wells (Fakher *et al.*, 2021). They can provide noticeable increases in a well's production if looked after and maintained within their optimum operating ranges. However, ESP failures are usually sudden and unanticipated and could become common occurrences. These failures cause significant financial impacts for the operators due to oil production deferments and high intervention costs (Takacs,

2009). In addition, excessive shutdowns and trips significantly reduce the run-life of an ESP (El Gindy *et al*., 2015).

An ESP system comprises both surface and downhole components (Takacs, 2009). The main downhole components consist of a multistage centrifugal pump, gas separator, electric motor, seal section, power cable, and downhole sensors. The main surface components are variable speed drive (VSD), transformers, surface electric cable, junction box, and wellhead. Figure 2-1 illustrates the standard setup of an ESP system.

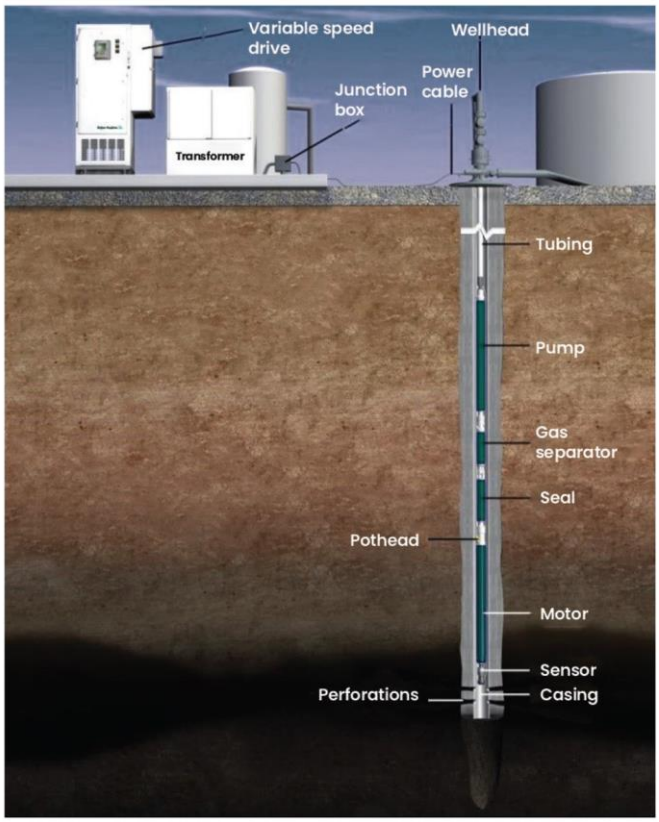

Figure 2-1: A Representative ESP System Schematic (Hughes, 2020)

The Electrical Submersible Pump system consists of multiple stages of centrifugal pumps stacked in series and coupled to a submersible electric motor. Variable speed drives (VSDs) enhance the ESP performance by altering the pump's operating frequency and, therefore, its speed.

The junction box connects the VSD cable to the downhole electrical cable and permits the venting of wellbore gas. The motor receives electrical power from the surface controls through the electrical cable, transforms it to mechanical energy, and transmits it to the pump impellers via a coupled shaft. Motor cooling is achieved as the produced fluid passes the motor housing. The shaft's mechanical energy is conveyed to the pump stages (Takacs, 2009).

Depending on the well's operational needs and completion design, the number of pump stages required for the desired flow rate may vary. Each stage is equipped with a rotating impeller and stationary diffusers. Figure 2-2 illustrates a multistage ESP system with a rotating impeller and a stationary diffuser at each stage. Figure 2-3 depicts a single-stage ESP pump with an impeller and a diffuser. By spinning the blades, the impeller promotes fluid flow and delivers kinetic energy to the fluids (Takacs, 2017). The fluid's kinetic energy is transformed to pressure potential at the diffuser. This process is carried out at each stage of the pump, as illustrated in Figure 2-4.

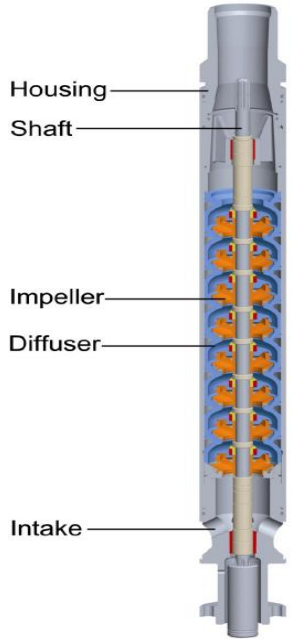

Figure 2-2: ESP Pump Cutaway (Hughes, 2020)

Figure 2-3: ESP Pump Stage-Impeller and Diffuser (Hughes, 2020)

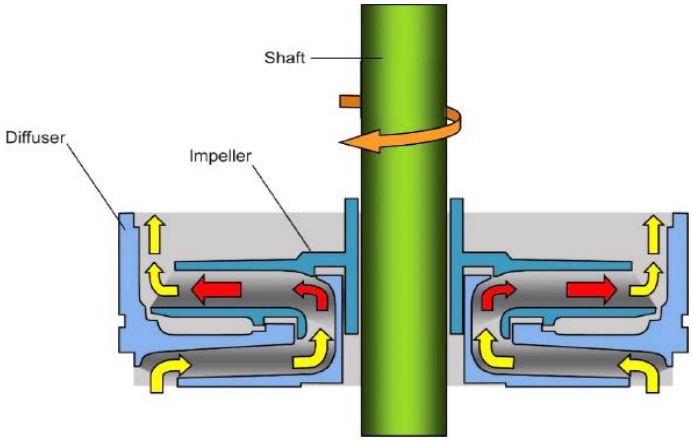

Figure 2-4: ESP Pump Stage-Impeller and Diffuser (Hughes, 2020)

The ESP seal section between the pump intake and the downhole electric motor provides many benefits to the system. It has thrust bearings that sustain the axial thrust generated by the pump (Hughes, 2020). In addition to isolating and protecting the motor from well fluids, the seal also equalizes the pressure in the wellbore and within the motor.

Free gas may be present in fluids produced from wells with low bottomhole pressures and high gas-oil ratios (GOR). Due to operational difficulties, such as cavitation or gas locking for gassy wells, ESPs may have a shorter run-life (Zhu and Zhang, 2018). Therefore, a gas separator is installed at the pump's intake in these wells to separate the free gas from the produced fluid.

Production monitoring and surveillance of ESP wells help extend their run life. Furthermore, an ESP monitoring system is highly beneficial to record a pump’s performance and

acquire valuable data of the downhole conditions. These data are used to analyze and predict failures before they occur. They involve a vast range of field conditions, including dynamic, static, and historical data. They are the key factors for data-driven models predicting ESP system failures (Abdelaziz *et al*., 2017).

2.1.1 ESP Advantages and Limitations

Utilizing ESP systems as an artificial lift technique has several operational advantages, including lower noise levels and a smaller surface footprint than other artificial lift methods. These systems may accommodate deviated or horizontal wells with doglegs of up to 10°/100 ft, but the pump must be installed in the straight section. ESPs excel at producing high liquid volumes (200-100,000 BPD) from moderate depths with a range of 1,000'-15,000' TVD (Fakher *et al*., 2021).

ESP operations provide several challenges. ESP run life may be significantly influenced under high sand and solids conditions, even though ESP systems can be built using specific abrasion-resistant materials (El Gindy *et al*., 2015). Production of highly viscous fluids (> 1,000 cp) and the presence of large amounts of free gas (> 5%) within the pump have negative impacts on performance (Barrios and Prado, 2009). Furthermore, precise well inflow data must be used in the design process, and the unit's capacity must match the well's deliverability. If not, expensive workover activities will be necessary to operate the pump (Williams *et al*., 2003).

2.2 Evolution of ESP Failure Prediction Over Time

Mohrbacher and Tabe (1984) discussed the ESP installation, operation, maintenance, and issues, including a system failure analysis. From the 1990s through 2010, several authors used statistical models to examine ESP problems. There were comparisons of failure distributions

among the system components (Higgs, 1994), different equipment types and models (Venkataraman, 1994), or different companies, platforms, and fields (Oliveira *et al*., 1997). In addition, historical trends have been used to assess the evolution of ESP run life through time. The authors Oliveira *et al*. (1997), Patterson (1993), and Venkataraman (1994) presented statistical distributions fitted to historical data to forecast future failure frequencies. Sawaryn *et al*. (1999) emphasized the need to include all the ESP system parts in the study. They suggested that the simulation time be extended to encompass the whole life of a field and appropriately measure ESP reliability. Furthermore, Alhanati *et al*. (2001) provided a standardized ESP failure analysis that included a list of all ESP-related failure types, failed items, failure descriptors, and root causes.

The failure rates of ESPs across an oilfield have been shown to vary significantly (Sawaryn *et al*., 2002). Given the number of ESPs installed in a given oilfield, ESP failure prediction research is mainly concentrated on computing population-level estimations and failure causes (Sawaryn and Ziegel, 2001). Sawaryn (2003) came up with analytic terms to describe failure patterns at a population level. Liu *et al*. (2010) also modified the data mining classification techniques for rod pump and tubing failures.

Despite significant academic research on ESP lifetime prediction, a few practitioners have employed predictive modelling in their works. Guo *et al*. created a binary classification model using support vector machine algorithm (SVM) to predict ESP failures using electrical and frequency data (Guo *et al*., 2015). Only the data from ESP sensors were utilized, with the most important characteristics being current, voltage, frequency, and duration of operation.

Gupta *et al*. (2016a) and Abdelaziz *et al*. (2017) used the principal component analysis (PCA) method as unsupervised machine learning model to detect patterns of stable and unstable ESP functioning. Following that, Andrade Marin *et al*., (2019) investigated various approaches

and opted to utilize a random forest as binary classification model as an example, which resulted in a high degree of accuracy and recall. Bhardwaj *et al*. (2019) used the principal component analysis (PCA) methodology and the gradient boosting technique (XGBoost) for the first time to discover anomalies. Subsequently, a short-term failure prediction using binary classification model was presented by Khabibullin *et al*. (2020) based on XGBoost, CatBoost, and LightGBM as the key models. Table 2-1 summarizes prior works on ESP Failure Prediction.

Table 2-1: Evolution of ESP Failure Prediction Over Time

Researchers	Dataset Size	ML Model Applied	Predicted Parameter
Guo *et al*., 2015	• High frequency data • No physical model used	Support Vector Machine (SVM)	• Prediction of anomalous operation • Failure modes
Gupta *et al.,* 2016a	• High frequency data*	Principal Component Analysis (PCA)	• Prediction of anomalous operation
Abdelaziz *et al.,* 2017	• High frequency data*	Principal Component Analysis (PCA)	• Prediction of anomalous operation
Andrade Marin *et al*., 2019	• 165 cases • High frequency data** • For 3 years of operation	Random Forest (RF)	• Long-term failure prediction (30 days).
Bhardwaj *et al.,* 2019	• 6 cases • 2 samples***	• PCA • XGBoost (extreme Gradient Boosting)	• Prediction of anomalous operation • Failure modes
Khabibullin *et al*., 2020	• 397 cases • For 4 years of operation • High frequency data are not available	• XGBoost • CatBoost • LightGBM	• Short-term failure prediction models (7 days)

*The data frequency is 1 per minute.
** The data frequency is 2 times per hour.
*** The data frequency is once a day and once a minute.

2.3 Evolution of ESP Monitoring System Over Time

The monitoring levels of ESP systems have changed over time. The first and simplest method of diagnosing an ESP system's failure is by using ammeter charts to measure and record the current drawn by the downhole motor (Takacs, 2009). The wellhead pressure is then utilized

to calculate the pump's head, followed by fluid shots to measure the fluid level in the casing. Fluid shots are useful but may be inaccurate owing to factors like foamy crude or completion configuration (Takacs, 2017). Recently, it is possible to safeguard an ESP by forcing the pump shutdown in case of underload or overload current values, which are signs of poor operating conditions.

The variable speed drive (VSD) was a major step in the growth of the ESP technology. This equipment gives operators an extra means of controlling an ESP’s functions, while also protecting pumps and motors from electrical stress. The main advantage of a VSD is that it allows the operator to modify the ESP's speed in response to variations in the well's productivity index or changing well conditions (Williams *et al.*, 2003).

The introduction of ESP Downhole Monitoring System was also a significant step in the evolution of ESP monitoring. This system gathers and transmits data from the downhole system to the surface, where it may be recorded and researched to optimize system performance and longevity. This helps to improve the accuracy of the most crucial production metrics. Downhole monitoring is becoming more prevalent, with a growing number of new installations being equipped with it (El Gindy *et al*., 2015). It analyzes well characteristics and delivers pump data, with the goal of improving ESP efficiency and reserve recovery rates (Sherif *et al*., 2019). It helps to keep track of some key operational parameters, as indicated in Table 2-2.

Table 2-2: Main Parameters Recorded in an ESP Downhole Monitoring System (Sherif *et al*., 2019; Gupta *et al*., 2016a)

Parameter	To determine:
Pump Intake Pressure (P_i), psi	Static and flowing bottom hole pressures
Pump Intake Temperature (T_i), °F	Gas Volume Fraction at pump intake
Pump Discharge Pressure (P_d), psi	Head and efficiency of the pump
Pump Discharge Temp (T_d), °F	Optional
Motor Temperature (T_m), °F	Operating temperature rise (Tm –Ti)
Current Leakage (mA)	Indication of impending ground fault conditions
Vibration (V_x & V_y)	Bearing mechanical wear and frequency that cause excessive vibration (Resonance).
Motor Frequency, Hz	Motor speed
Motor Current (I_m), Amps	Changes in fluid density and power consumption

The aforementioned parameters indicate the most relevant aspects to consider when analysing ESP failures (Adesanwo *et al*., 2016; Bermudez *et al*., 2014; Grassian *et al*., 2017). Even though appropriate monitoring helps delay system failures, they may still happen for a variety of reasons (Takacs, 2017). Because the electrical system is typically the weakest link in an ESP system, most failures are electrical, usually resulting from a mechanical problem as the underlying cause of the failure (Pennel *et al.*, 2018). Each failure must be fully probed, and its fundamental reason must be determined. Mubarak *et al*. (2003) researched ESP failures in Wafra field over a four-year period and discovered that the most common causes of ESP failures include motor failures (40%), pump failures (22%), cable failure (26%), and others (12%). Furthermore, according to Al-Sadah (2014), the most common reasons for ESP failures are burnt motor (28%), damaged cable/MLE (27%), and damaged penetrator (24%) based on the Dismantle, Inspection and Failure Analysis (DIFA) results. In Chapter 3, a comprehensive set of data from multiple oil wells with ESP systems is used to further analyse the ESP failures over time and explain the root physical causes.

2.4 ESP Performance and Gas Interference

This section discusses the methods used to analyze the gas interference in the ESP system. The pump performance evaluation and efficiency calculations will also be discussed here.

2.4.1 Gas Interference in ESP Systems

Free gas has a significant effect on the performance of ESP's. Earlier research has shown that the gas involvement reduces the ESP's hydraulic head. Unlike the diffuser, the performance of an ESP impeller is strongly influenced by the free gas entering the pump. The presence of free gas in the impeller lowers the effective area available for liquid flow, increases hydraulic losses in the stage, and decreases the pump's head produced (Takacs, 2017). In gaseous conditions, ESP flow behaviors such as surging, and the development of gas pockets worsen the ESP pressure boosting. Vibration and limited-service life might arise from surging, and gas pockets can significantly restrict liquid production rates (Zhu and Zhang, 2018).

A typical pump performance curve is seen in Figure 2-5. Barrios and Prado (2009) and Zhu and Zhang (2014), among many others, investigated the influence of free gas on the impeller performance using experiments and Computational Fluid Dynamics (CFD) simulations. According to them, pump performance was not significantly changed when the free-gas phase was evenly dispersed in the liquid. The free gas increases the total volume of fluid that the pump must handle. As a result, the pump's operating range moves to the right on its performance curve, decreasing its head (Barrios and Prado, 2009).

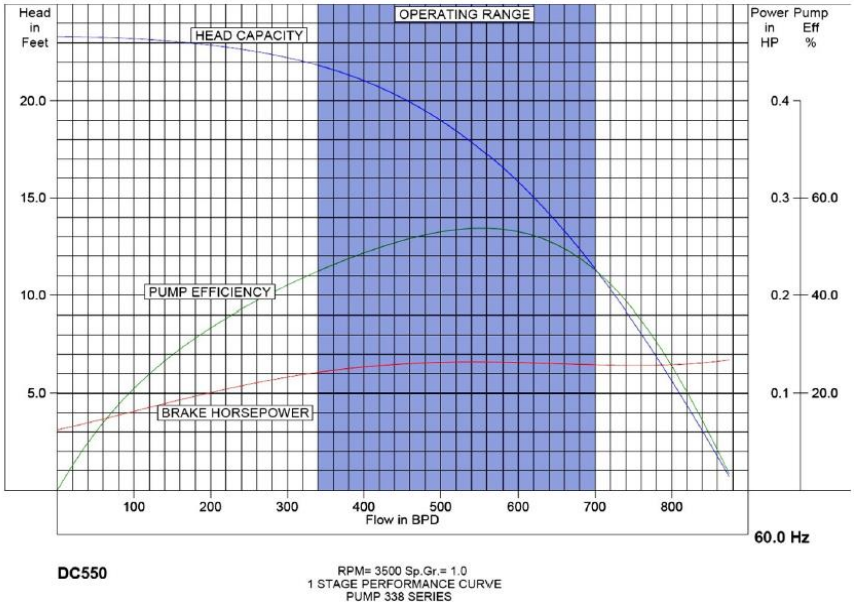

Figure 2-5: Typical ESP Performance Curve (Hughes, 2020)

The impeller flow is heterogeneous with gas, because of the considerable density difference between the two phases (Takacs, 2017). Since liquid particles are significantly denser than gas particles, they have higher kinetic energy and flow faster through the vane openings of the impeller. As a result, centrifugal forces overcome the turbulence and separate the gas and liquid particles at the vane opening (Takacs, 2017), resulting in gas bubble coalescence.

The low-pressure sides of the impeller vanes are more favorable to gas flow, while the high-pressure sides are more favorable to liquid flow. When modest amounts of gas reach the pump, the liquid flow drags little gas bubbles toward the diffuser without phase slip (Zhu and Zhang, 2018). Bubbles in the impeller intake grow due to increased free gas at the pump suction and the accumulation of smaller bubbles already present. Gas is trapped at the impeller intake when bubbles reach a critical value, resulting in further bubble generation and the creation of a gas pocket (Barrios and Prado, 2009). A gas lock may occur if these pockets become large enough to obstruct the liquid flow through the impeller. As a result, the ESP system is put at risk, as shown in Figure 2-6.

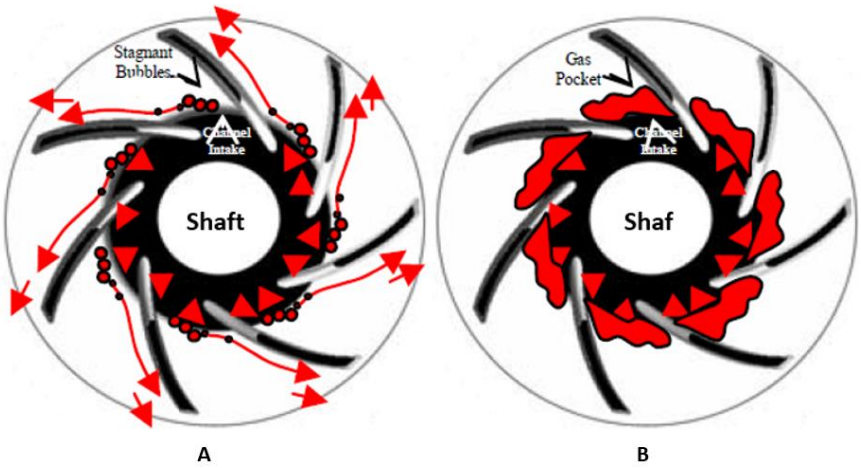

Figure 2-6: After Barrios and Prado (2009); A) Stagnant Bubbles at Channel Intake. B) Gas Pockets at Channel Intake Causing Surging

2.4.2 ESP Performance Evaluation

In 1986, Turpin *et al*. established the first and a commonly used model for estimating ESP performance in the presence of gas. They employed the Turpin correlation parameter (ϕ) to determine ESP operating conditions based on gas volumetric fraction and pump intake pressure. When ($\phi > 1$), the ESP surge condition is formed. This approach is often used to assess an ESP's stable operating limitations. As shown in Figure 2-7, the ESP is stable below the Turpin curve, but unstable above the curve due to the possibility of gas interference (Turpin *et al*., 1986).

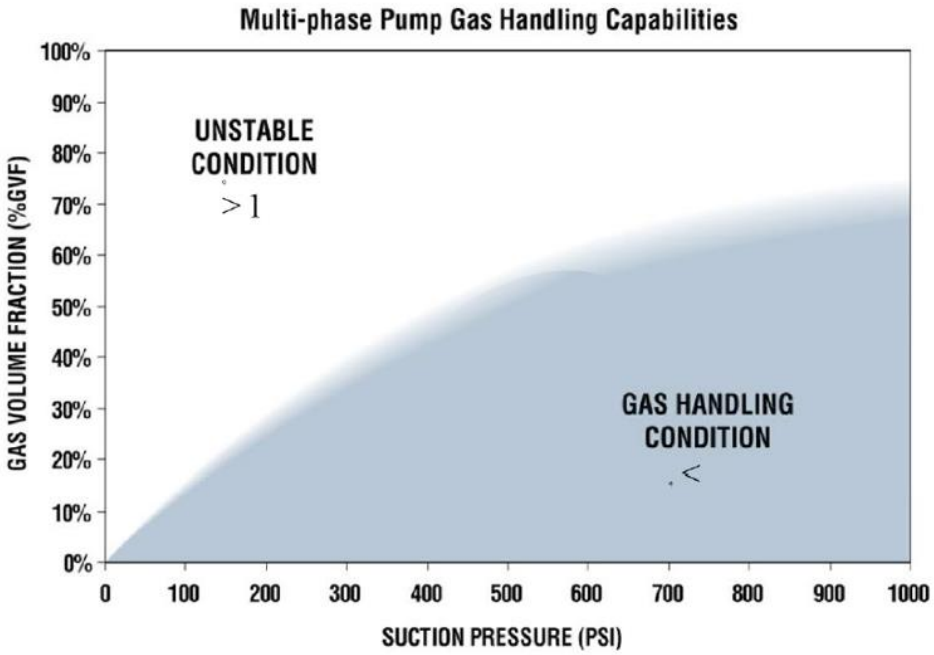

Figure 2-7: Turpin Curve (Hughes, 2020)

Later, Dunbar (1989) proposed an empirical correlation, agreeing with Turpin that the pump intake pressure impacts gas volumetric fraction (GVF) estimation. Cirilo (1998) devised an empirical model to calculate the critical GVF for the ESP surge initiation stage. In 2003, Duran and Prado created an empirical model to predict stage pressure increments of bubble flow. They proposed that the surging correlation be extended to include gas and liquid densities as well. This correlation is more often employed in pump design than in monitoring.

The multiphase impact on ESP performance has recently been examined by Zhu *et al*. (2019). ESP gas-liquid flow patterns were analyzed using a novel mechanistic model that correlates critical gas fraction, angular speed, surface tension, liquid flow rate, impeller volume, and rotor radius. This approach is mostly concerned with ESP design. According to prior research, pump intake pressure strongly impacts the GVF estimation in the ESP. The surging correlations in the literature are summarized in Table 2-3.

Table 2-3: Surging Correlations in Literature

Study	Correlation	Equation
Turpin *et al*., 1986	$\Phi = \frac{2000\ q_s}{3\ P_i Q}$ $\lambda_c = \frac{q_s}{q_s + Q}$ Where: • Φ is Turpin correlation parameter (psi^{-1}) • q_s is gas volumetric flow rate (BPD) at pump intake • Q is liquid volumetric flow rate (BPD) at pump intake • P_i is pump intake pressure (psi) • λ_c is gas volumetric fraction, GVF	**2-1** **2-2**
Cirilo, 1998	$\lambda_c = 0.0187\ {P_i}^{0.4342}$	**2-3**
Duran and Prado, 2003	$\frac{Q_g}{Q_{max}} = \left(5.580 \times \frac{\rho_g}{\rho_l} + 0.098\right)\left(\frac{Q_l}{Q_{max}}\right)^{1.421}$ Where: • $\frac{Q_g}{Q_{max}}$ and $\frac{Q_l}{Q_{max}}$ are normalized gas and liquid rates (estimated at surging initiation) • $Q_{max}$is maximum pump rate	**2-4**
Zhu *et al*., 2019	$\lambda_G = \frac{d_{max}}{10.056\left(\frac{\sigma}{\rho_c}\right)^{0.6}\left(\frac{\Delta P_{pump} Q_l}{\rho_c V}\right)^{-0.4}\left(\frac{\rho_c}{\rho_d}\right)^{0.2}}$ Where: • λ_G is the gas volumetric fraction, GVF • σ is surface tension (N/m) • V is impeller volume (m^3) • ΔP_{pump} is pressure increment of the single-stage ESP • Subscripts c and d denote the continuous and dispersed phases. • d_{max} is maximum bubble size in turbulent flow field $\lambda_{cr} = \frac{2\left[\frac{0.4\ \sigma}{\left(\rho_l - \rho_g\right)\Omega^2 R_1}\right]^{0.5}}{10.056\left(\frac{\sigma}{\rho_l}\right)^{0.6}\left(\frac{\Delta P_{pump} Q_l}{\rho_l V}\right)^{-0.4}\left(\frac{\rho_l}{\rho_g}\right)^{0.2}}$ Where: • λ_{cr} is the critical GVF at which ESP surging initiates • Ω is Angular speed (rad/s) • $R_1$is rotor radius (m) • The numerator represents the critical bubble diameter in multiphase flow	**2-5** **2-6**

2.5 Literature Review Summary

This chapter has presented a detailed review of the available literature on this study's primary topic. While providing a comprehensive overview of ESP systems' key advantages and

limitations, the discussion focuses on how ESP failure prediction and system monitoring have progressed through time and explores the gas interference in the ESP system further.

Over time, the ESP monitoring techniques have evolved, leading to the development of the ESP downhole monitoring systems. These systems are essential for real-time monitoring to meet operational needs, delivering reliable data on pump and reservoir performance to improve ESP system performance and longevity. Furthermore, ESP failure prediction has developed through time, particularly in the last few years. As seen in Table 2-1, most prior studies used unsupervised ML models to predict anomalies within the normal ESP operating time. In contrast, the remainder employed supervised binary classification to predict an ESP failure with limited parameters. Prior research has mostly focused on detecting ESP failure and has neglected to investigate and predict the root cause of ESP failures. Predicting the fundamental cause of ESP failure is essential to choose the most effective course of action to mitigate it. It is however nearly impossible to physically model the performance of a complete ESP setup, considering the multitude of electrical, mechanical, and fluid-related components coupled to form an integrated ESP system.

This study proposes a novel approach utilizing supervised multi-class classification models to construct a predictive tool for ESP specific failure modes (SFMs) utilizing a comprehensive dataset and various machine learning algorithms. Also, this study investigates the gas interference in the ESP system and the pump efficiency based on the field data. The proposed ML models incorporate the physical aspects of ESP operations to provide a hybrid approach for failure prediction. The best model with the highest accuracy in forecasting ESP failures is identified. In addition, several prediction periods are compared in terms of the model performance.

Chapter 3 Statistical Analysis

This chapter describes the statistical analysis and a detailed ESP failure overview utilizing a set of field data on ESP failures. In addition, two case studies from Kuwait Oil Company assets are discussed in depth to explore the failure types and their implications on the ESP and well production. First, the available data are detailed.

3.1 Operational Field Data Overview

This research is based on actual field data from Kuwait Oil Company (KOC) assets. For the evaluation, these data are categorized into static, dynamic, and historical data from over five years of production of 10 active wells. The dynamic high-frequency data (per minute) were obtained from surface and downhole ESP monitoring equipment (VSD, pump, and wellhead). The static data include the well completion, ESP design, and reservoir fluid information. Historical operational data are used to supplement the analysis. The data were collected from wells with high ESP failure rates to investigate the general ESP failures and their specific failure modes.

Large amounts of high-frequency ESP operational data (~24.4 MM data points) were acquired to test and validate the approach and assess the robustness of various methods in this research. The dataset includes 10 essential ESP information variables to create analytical models and predict failures. The ESP data variables supplied by the VSD, and the well are shown in Table 3-1.

Table 3-1: ESP Data Variables

	Variables	Description
1	*WHP [psi]*	Wellhead pressure, psi
2	*CHP [psi]*	Casing head pressure, psi
3	*FLP [psi]*	Flowline pressure, psi
4	*I_MOTOR*	Motor current, Amps
5	*V_MOTOR*	Motor voltage, Volts
6	*MOT_FREQ*	Motor frequency, Hz
7	P_i *[psi]*	Pump intake pressure, psi
8	P_d *[psi]*	Pump discharge pressure, psi
9	T_i *[Deg F]*	Pump suction temperature, °F
10	T_m *[Deg F]*	Motor temperature, °F

All of the wells in this research had identical ESP systems setup, including a multistage centrifugal pump, gas separator, electric motor, seal section, power cable, downhole sensors, VSD, transformers, surface electric cable, junction box, and wellhead. The number of ESP pump stages required to achieve the target flow rate vary based on the well's operating requirements and completion design. Moreover, both radial and mixed flow ESP pumps are included. The smaller flow pumps often have a radial design, while the larger flow pumps typically have a mixed design.

The ESP motors used in this study are typically two-pole, three-phase induction motors with squirrel cages. The motors operate at a nominal speed of 3500 rpm at 60 Hz or 2915 rpm at 50 Hz due to their two-pole configuration. The motors operate on three-phase power between 230 and 5,000 volts and between 12 and 200 amps. The diameter and length of the motors determine their horsepower ratings, and since there are no wires running along their length, the motors may be built somewhat larger than the pumps. These wells are vertical, shallow (TVD of 4,500-5,500 ft), and producing medium oil (27-31°API). The ranges of physical parameters of these wells are shown in Table 3-2.

Table 3-2:Well's Fluid Characteristics

Fluid Characteristics	Range	
	From	To
Stock tank Oil Gravity, °API	27	31.5
Gas Oil Ratio, GOR (SCF/STB)	425	900
Initial Dissolved Gas-Oil-Ratio, R_{si} (SCF/STB)	425	450
Formation Volume Factor at bubble point, B_{ob} (RB/STB)	1.22	1.23
Reservoir Temperature, T_{res} (°F)	131	141
Bubble Point Pressure, P_b (psi)	1478	1690
Gas Specific Gravity, γ_g	0.81	0.95
Gas Formation Volume Factor, B_g (ft^3/SCF)	0.005888	0.006928

3.2 Statistical Failure Data Evaluation

The findings of the statistical analysis are depicted Figure 3-1, after dividing the failures into a few main categories. In addition, Figure 3-2 shows a more specific division of the most statistically important causes for ESP failures as labeled by the field engineers. The detailed specific failures are depicted and described for electrical failures, motor failures, gas effects, and other failures in Table 3-3, Table 3-4, Table 3-5, and Table 3-6, respectively. Some of the most witnessed general failure roots are:

- Electrical Failure (61%): including power fail, under-voltage, over-voltage, phase unbalance, and overcurrent.
- Motor Failure (18%): including motor voltage unbalance.
- Gas Effect (13%): excessive gas content locked in the pump causing the current to drop rapidly because of motor underload.

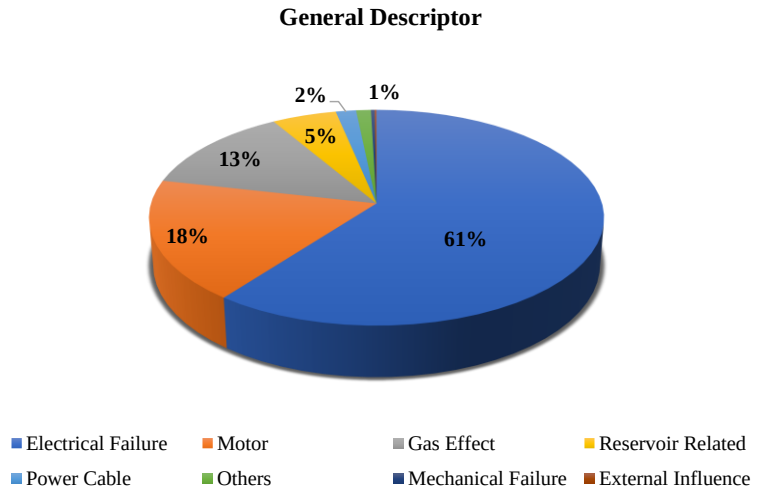

Figure 3-1: ESP Failure Statistical Analysis-General Descriptor

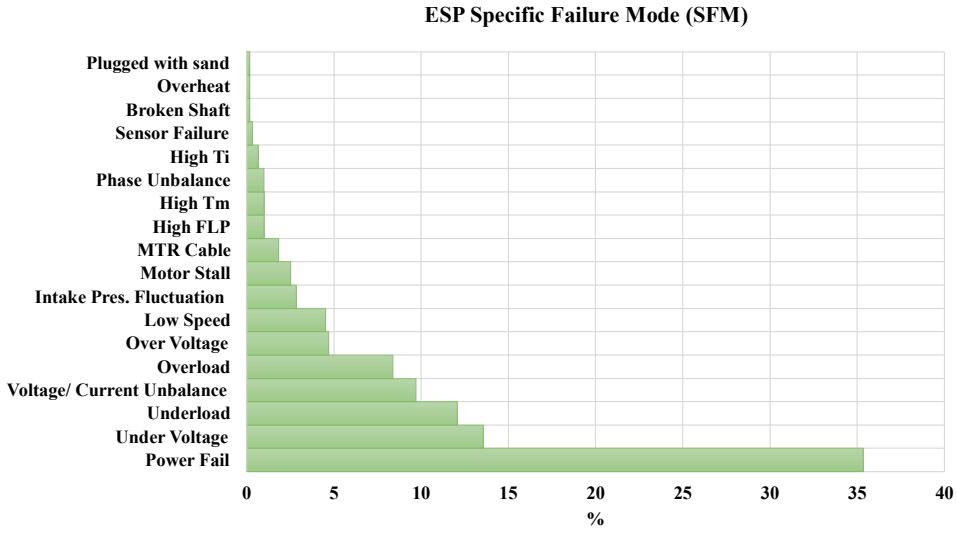

Figure 3-2: ESP Failure Statistical Analysis-Specific Failure Mode

Table 3-3: The Likely Specific Causes of ESP Electrical Failures

Specific Failure Mode	Likely Cause
Power Fail	Electrical issues with the power supply.
Under-Voltage	Intermittent under voltages and over voltages can be caused by a poorly controlled primary power supply or wiring problems.
Over-Voltage	
Phase Unbalance	Unbalanced power input
Overcurrent	Over/undervoltage
	Excess flow
	Bad bearings
	Foreign material in the pump
	Jammed pump

Table 3-4: The Likely Specific Causes of ESP Motor Failures

Specific Failure Mode	Likely Cause
Voltage/ Current Unbalance	Unbalanced power input resulting in overheating and reduced motor life
Motor Overload	Loads exceed the motor design
	Mechanical failure of the downhole equipment
	An increase in the density of the fluid being pumped.
	A change in the supply voltage
	Debris in the pump causing impeller drag (locked rotor)
Low Speed	High torque
Motor Stall	The motor's speed will automatically decrease if the load exceeds the motor's capacity, at which point the motor will slow down or stop spinning.
Overheat	Motor's rotor overheating due to improper cooling
High T_m	High Motor Winding Temperature

Table 3-5: The Likely Specific Causes of ESP Gas Effect Failures

Specific Failure Mode	Likely Cause
Motor Undercurrent/ Underload	Restricted flow
	Lack of liquid to pump (reduced load on the motor).
	Broken shaft coupling or belt
	Gas locking causes underload shutdown
Intake pressure fluctuation	High gas production from the reservoir

Table 3-6: The Descriptions of the Other ESP Specific Failures

Specific Failure Mode	Description
Broken Shaft	A broken ESP shaft is categorized as a mechanical failure
High T_i	High pump suction temperature due to the high fluid temperature entering the pump
MTR Cable	A power cable failure caused by high temperature
Plugged with sand	External substances (sand) interference with the ESP assembly
Sensor Failure	Downhole sensor failure results in no sensor readings at the surface.

3.3 ESP Specific Failures Overview

In this study, several distinct ESP failures are analysed with regards to the early warning signs prior to the actual failures. The trends in the data are depicted with time, starting from two weeks before each specific failure mode (SFM). Common warning signs are assessed using ten

samples from each SFM presented in this section. The objective is to have a general understanding of the data trends associated with each SFM.

3.3.1 SFMs Associated with Electrical Failures:

ESP electrical failures are prevalent and have been attributed with both surface and downhole issues (Alhanati *et al.*, 2001). Increased downhole pressure and temperature often induce surface system failures by straining the electrical supply, resulting in failure of electrical components (Fakher *et al.*, 2021). Moreover, the downhole system fails if any of the electrical components in the ESP assembly fails, including the power cable, motor (i.e., stator), or the downhole sensor. As shown earlier, electrical failures account for 61% of all ESP operational failures. The trends in the acquired data prior to some of the ESP electrical failures are shown in this section.

i. <u>Under-Voltage:</u>

The under-voltage specific failure is caused by a main power supply with inadequate control or a wiring problem. This SFM has the potential to cause the pump motor to overheat, resulting in motor failure. Multiple ESP variables were explored to identify the patterns prior to the actual failure. Figure 3-3 illustrates a sample of ten under-voltage ESP specific failure occurrences from five wells, illustrating failure impacts on motor current prior to the actual failures. These figures illustrate two under-voltage failures for each of the five wells at separate intervals throughout the operating life of the wells (the second failure is designated by "*"), all of which are preceded by an increase in motor current. For under-voltage, the average time before failure (TBF) when motor current exhibits early warning signs is 1.4 days.

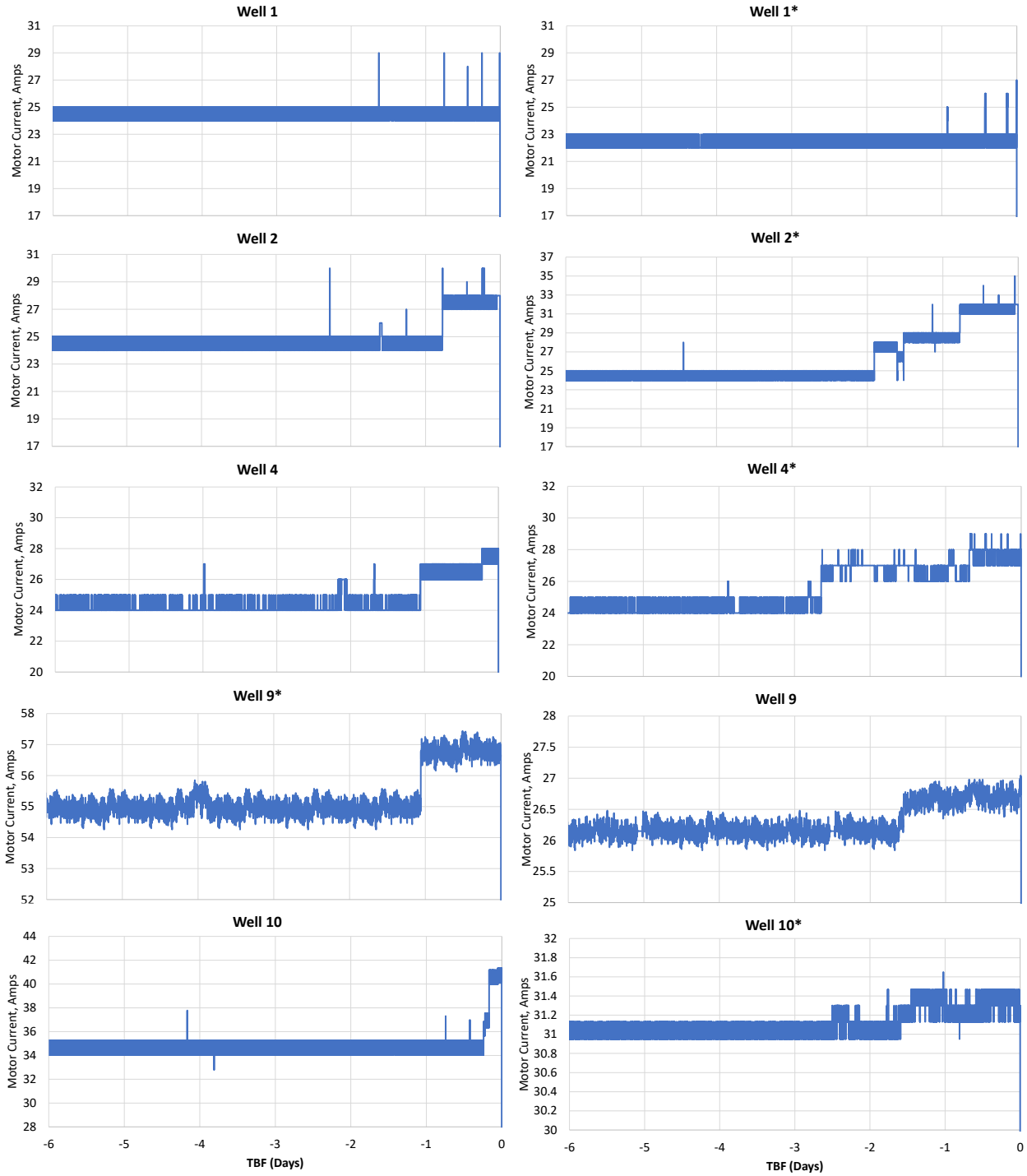

Figure 3-3: Under-voltage Specific Failure Effects on Motor Current

From this point on, only two of the analyzed cases will be presented for each failure, for the purpose of brevity. Figure 3-4 depicts motor voltage and motor current data for two weeks

prior to two examples of under-voltage failures in Wells 2 and 4 of our datasets. The x-axis indicates the time before failure (TBF) in days. Both voltage and current diverge from their typical trends at almost the same time in both wells. The voltage of the motor in Well 2 begins to drop 0.8 days prior to failure and continues to decline until the actual failure. The motor current rises almost in synch with the voltage decline until the failure. For Well 4, a reduction in voltage is first seen 2.1 days before failure, while the motor current increases. As the voltage of the motor drops, the motor draws greater current, which causes the motor to overheat, resulting in sudden failure. These wells display early warning indicators of a condition requiring immediate attention.

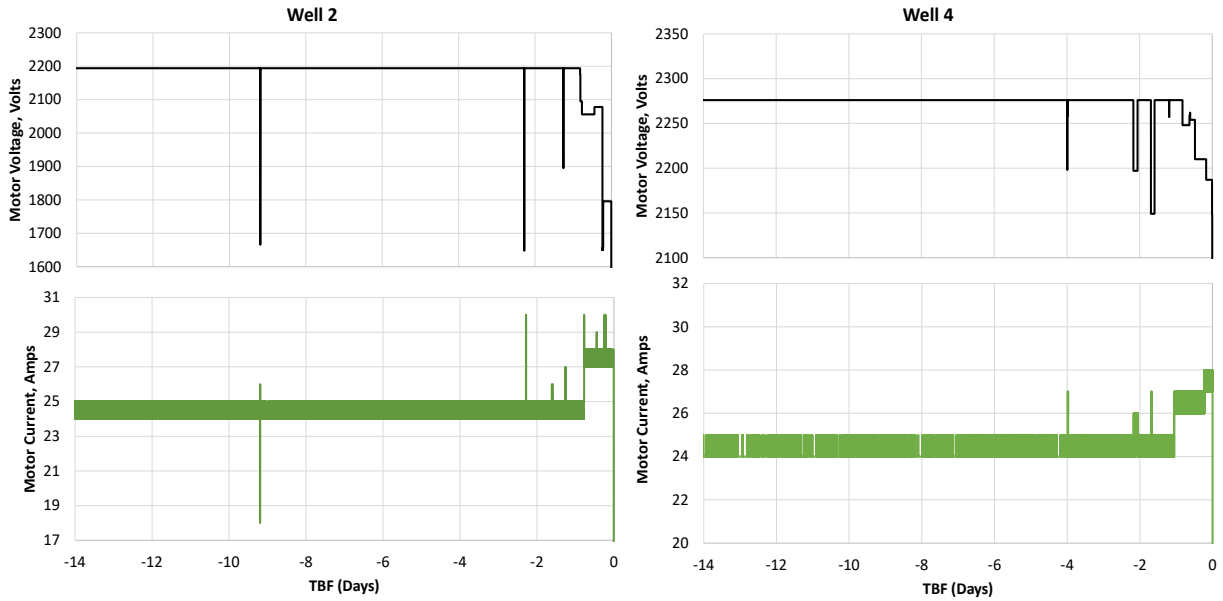

Figure 3-4: Under-voltage Specific Failure Effects on Motor Current and Voltage

ii. Power Fail:

A power fail occurs due to unbalanced phases, voltage spikes, presence of harmonics (distorted current and voltage), or lightning strikes. Consequently, the ESP motor and power cable are overheated. Takacs (2017) recommended a clean electric supply or a VSD unit with sinusoidal output to alleviate these issues.

Figure **3-5** depicts examples of motor voltage, motor current, and wellhead pressure for two weeks prior to two power failures in Wells 3 and 6. The figure shows the declining trend of wellhead pressure within a day before failure for both wells. In addition, voltage spikes (surges) are seen in Well 3 beginning 0.78 days before failure. Similar trends are found for the Well 6 case, where the spikes start 2.9 days before the failure. Voltage surge is the rapid increase in voltage over a very brief period in a power system (Takacs, 2009). In addition, the motor current in Well 3 exhibits an increasing trend three days before failure. Similar observation is made for Well 6, starting 2.8 days before the failure, exhibiting early warning signs of a problem that demands attention.

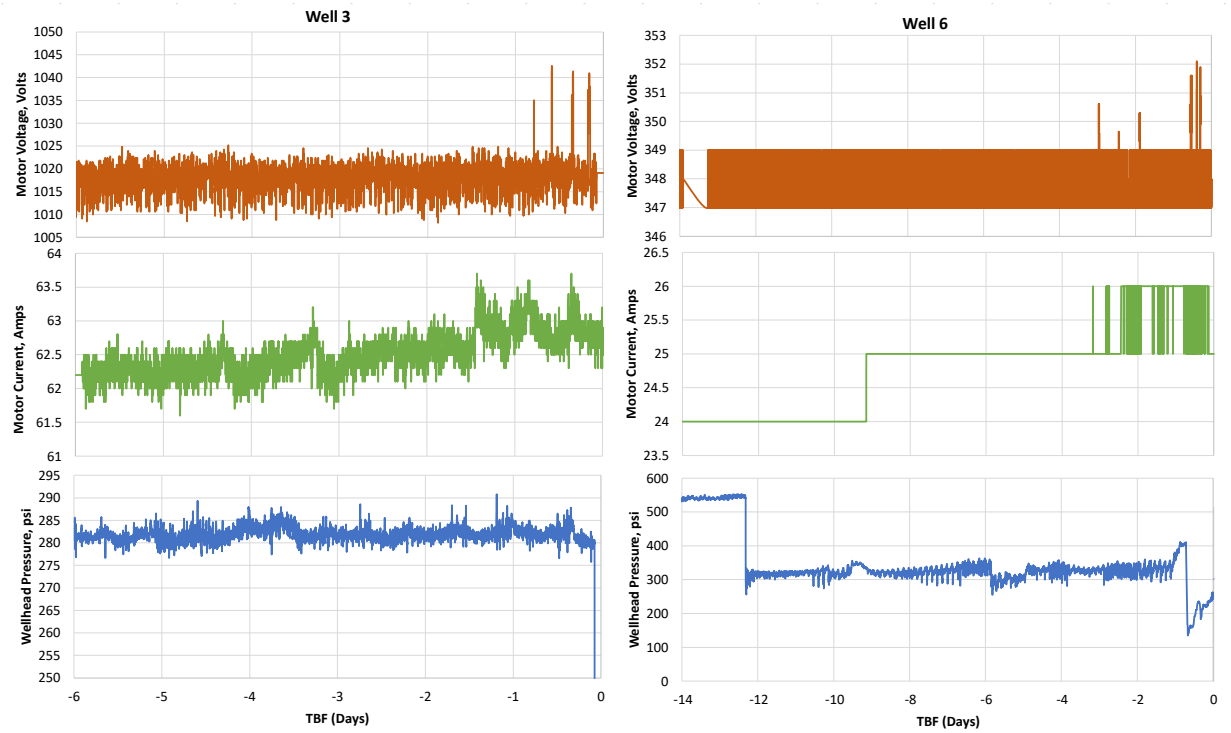

Figure 3-5: Power Fail Specific Failure Effects on Motor Voltage, Current and Wellhead Pressure

3.3.2 SFMs Associated with Motor Failures:

Most ESP motor failures are due to electrical issues (Alhanati *et al*., 2001). Based on the investigated field data, voltage/current unbalanced, motor overload, and high motor winding temperature are some of the motor related failures.

The ESP motors utilized in this study are generally two-pole, three-phase AC induction motors, as detailed in Section 3.1. Since a three-phase power system is designed to function with the phases (Lines) balanced, phase imbalance occurs if one or more of the line-to-line voltages in the system are mismatched (Hughes, 2020). High current unbalance is caused by voltage unbalance at the motor terminals, which in turn increases vibration, causes the motor to overheat, and shortens its lifespan. An unstable utility supply, an improperly balanced transformer, an uneven distribution of single-phase loads on the same electrical system, or faulty wiring might all contribute to phase unbalance (Takacs, 2017). As shown earlier, motor failures account for 18% of all ESP operational failures. The trends in the acquired data prior to some of the ESP motor failures are shown in this section.

i. <u>Motor Overload:</u>

An overload specific failure happens when the motor draws an excessive amount of current, resulting in excessive power consumption. As a result, the motor may overheat, leading to motor damage (Alhanati *et al*., 2001). There are various factors that may contribute to an overload, including an improperly sized motor, an unexpectedly high fluid specific gravity raising the Total Dynamic Head (TDH) over the design value, and inconsistent motor voltage (Sawaryn *et al*., 2002). Figure 3-6 depicts wellhead pressure, motor current and motor temperature trends for two weeks prior to two cases of overload failures in Wells 3 and 10. The x-axis indicates the time before failure (TBF) in days. The figure shows the declining trend of wellhead pressure as failure

time approaches for both wells. Moreover, the motor current in Well 3 suddenly increases 3 days before failure. The trend is similar for Well 10, where the deviation begins 1.8 days before the failure. The motor temperature also rises as the failure approaches, beginning 2.7 days in advance for Well 10 case and 2.1 days before for Well 3. These wells may exhibit early warning signs of a problem that demands attention.

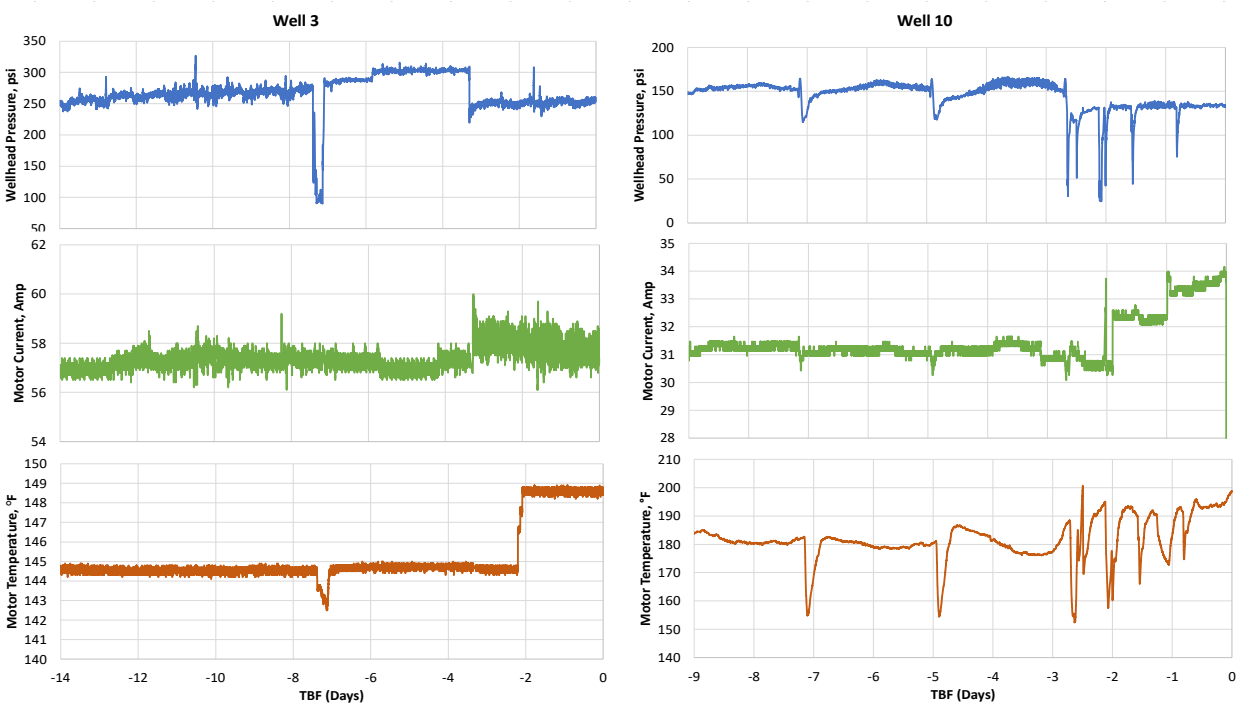

Figure 3-6: Motor Overload Specific Failure

ii. <u>High Motor Temperature:</u>

A high motor temperature happens when excessive voltage supply or drawn motor current lead to overheating issues (Takacs, 2009). Overheating is a primary cause of motor failure, particularly when the motor is forced to work harder or is placed under an unexpected load. As the speed of the motor and, by extension, the rotation of the pump's shaft rises, the moving components get overheated due to the increased friction (Betonico *et al.*, 2015). Figure 3-7 displays motor voltage, motor current, and motor temperature prior to two cases of motor temperature failure in

Wells 8 and 1. The x-axis represents the TBF in days. The graph depicts the increasing trend of motor current as the failure time for both wells' approaches. In addition, the motor voltage in Well 8 surges rapidly 0.6 days before the failure. Similar observation is made for Well 1, where the deviation starts one day before the failure. As the failure nears, the motor temperature also begins to increase, starting 0.8 days before failures in both wells. These wells display early warning signals of a serious situation.

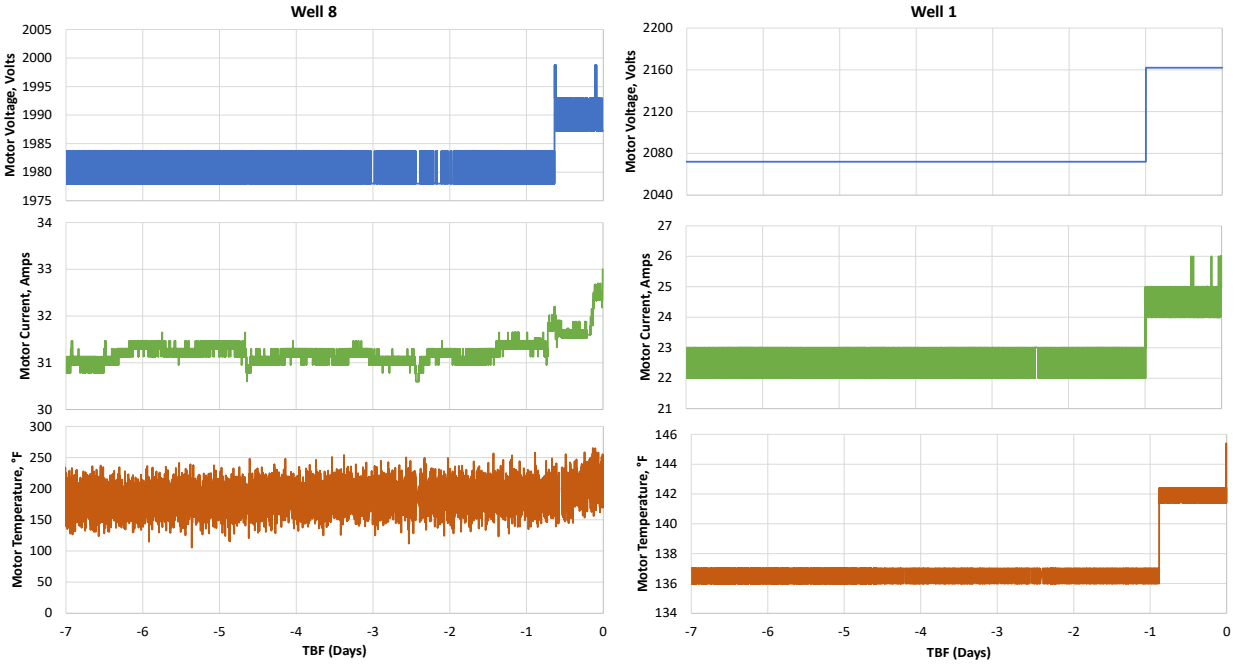

Figure 3-7: High Motor Temperature Specific Failure

3.3.3 Effects of Other Failures:

i. Motor Underload:

An underload failure often indicates pumped off or gas locked conditions (Takacs, 2017). Different high frequency ESP variables were investigated to observe the patterns prior to the failure. Figure 3-8 depicts pump intake pressure and motor current during the two weeks prior to two examples of underload failures in Well 8 and Well 9. The x-axis indicates the time before

failure (TBF). The figure shows the rising trend of pump intake pressure as failure time approaches for both wells. Moreover, the motor current in Well 8 deviates from its regular pattern at 6.3 days, with a decreasing trend as the failure approaches. Similar observation is made for Well 9, where the deviation begins 1.8 days prior to the failure.

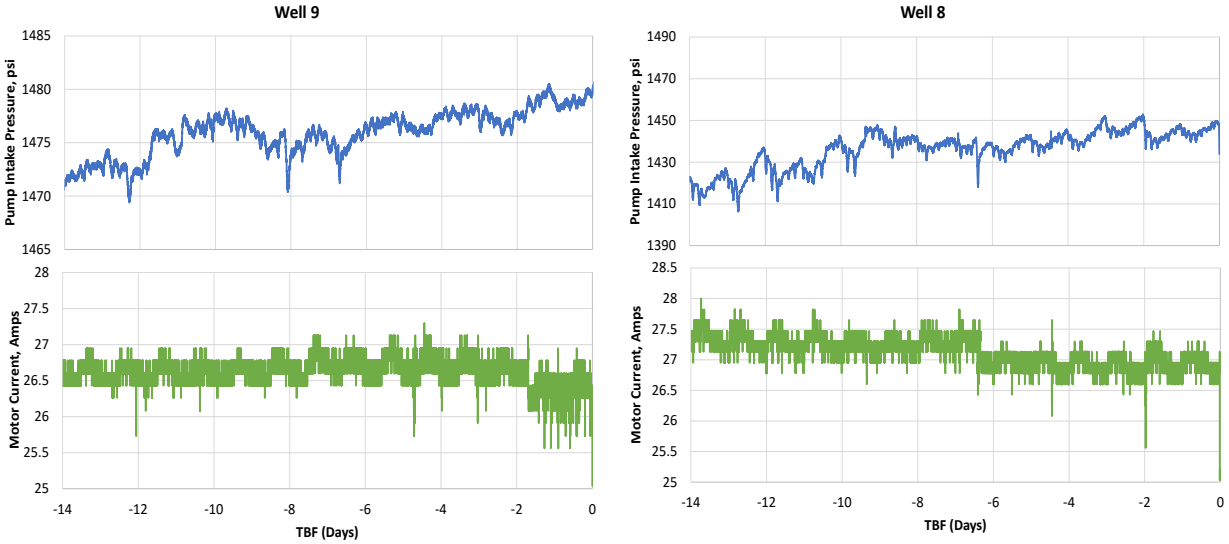

Figure 3-8: Underload Specific Failure

ii. <u>High Cable Temperature:</u>

High cable temperature indicates a power cable failure caused by high temperature. All downhole components, including the motor and sensors, are powered via an electrical cable that connects to a source of electricity at the surface. Instability in the power supply or a sudden increase in power demands might overload the cable. This abrupt spike may be caused by gas phase rise causing motor rotor speed to increase, which overloads the ESP (Takacs, 2009). Several other factors may also contribute to cable failures including mechanical damage (crushing, cutting, etc.) during running or pulling operations, cable corrosion, insulation deterioration due to high temperature or wellbore gases, and larger than design currents raising the cable temperature and inducing insulation failure (Takacs, 2017).

Figure 3-9 displays motor voltage and motor current during the two weeks prior to two cases of high temperature cable failure in Wells 2 and 9. The x-axis represents the time before failure (TBF). The graph depicts a sharp increase in motor current around one day before the failure for both wells. In addition, the motor voltage deviates from its normal pattern and increases with many spikes 1.2 days ahead of the failure in Well 2 and one day prior to the failure in Well 9.

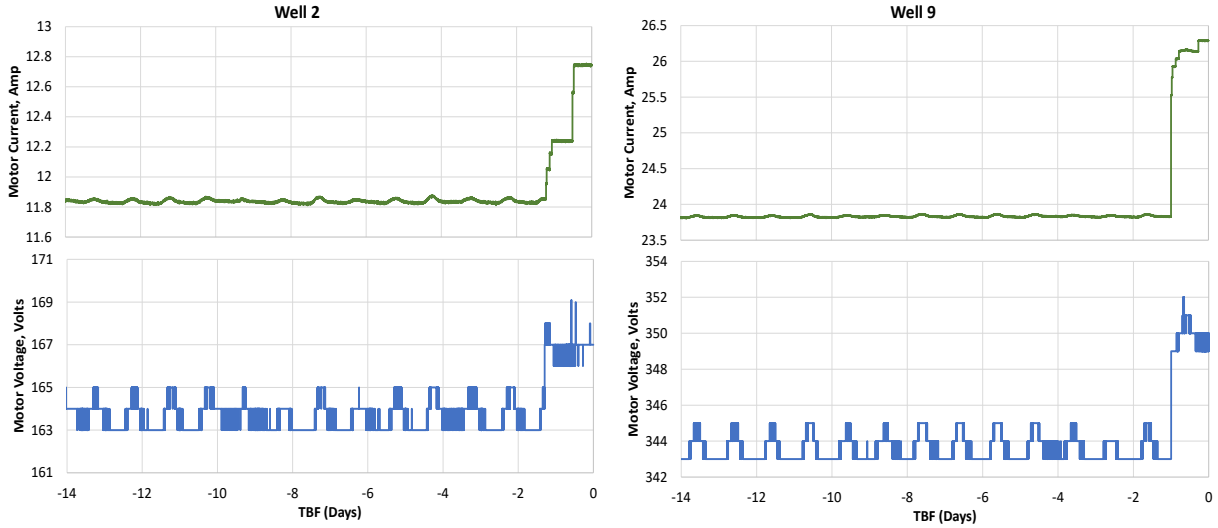

Figure 3-9: High MTR Temperature Specific Failure

iii. Sand Interference:

Plugging with sand is a specific failure caused by aggressive production of sand or other abrasive solid materials along with the well fluid. It is common for the production fluid to include sand, water, and gas. In extreme circumstances, however, the accumulation of solids in the pump's intake causes the pump to fail due to a full blockage of the intake (El Gindy *et al.*, 2015). One of the most common reasons of ESP failure, wear, and performance degradation is abrasive damage to the pump assembly due to sand interference (Williams *et al.*, 2003). Several factors, such as the degree of cementation and consolidation, high production rate, and high fluid viscosity (higher drag force; more sand) impact sand production (Takacs, 2009).

In sandy wells, system shutdown causes the solids to descend and settle on top of the pump, causing damage to the ESP's internal components. A hard/ rock start happens when an ESP is restarted with substantial quantities of sand laying on top of it, causing wear, reduced efficiency, and shorter run life. A broken shaft, motor burnout or an electrical failure owing to excessive current may come as results of this specific failure (Williams *et al.*, 2003).

Figure 3-10 depicts motor current during the two weeks prior to two cases of sand failure events in Wells 10 and 2. The x-axis indicates TBF in days. The graph illustrates the behavior of the motor current during multiple brief shutdowns and hard starts in Well 10. For Well 2, similar observations are made 1.7 days before the failure. These wells exhibit early warning signs of severe sand interference, which might lead to shaft failure with time.

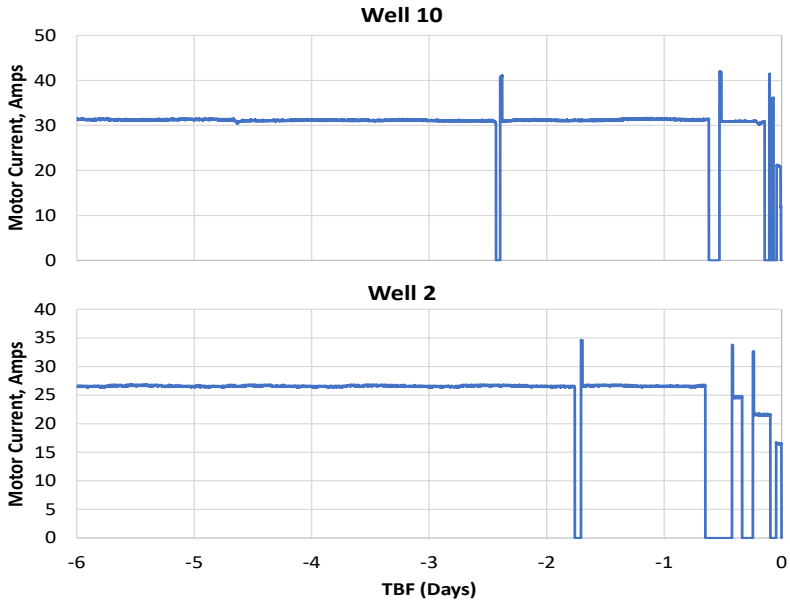

Figure 3-10: Plugged with Sand Specific Failure

iv. Downhole Sensor Failure:

Downhole ESP sensors monitor a wide range of parameters and transfer the data to the surface in real time. These sensors may record downhole pressures, temperatures, flowrate, system

vibration, and the system's electrical integrity (El Gindy *et al*., 2015). A sensor failure refers to a loss of communication with the ESP downhole assembly, which may be disastrous. Any electrical component failure in the ESP assembly, high temperature, hydrogen sulfide and carbonic acid corrosion may cause downhole sensors to fail (Medina *et al*., 2012). To mitigate sensor failure or malfunction, sensors must be designed to resist harsh downhole conditions (Medina et al., 2012).

Figure 3-11 illustrates motor voltage during the two weeks prior to two sensor failure occurrences at Wells 5 and 6. The graph depicts motor voltage spikes worsening with time, beginning 2.8 days before the failure for Well 5 and 1.2 days before the failure in Well 6. For this SFM, the only variable with early symptoms is the motor voltage.

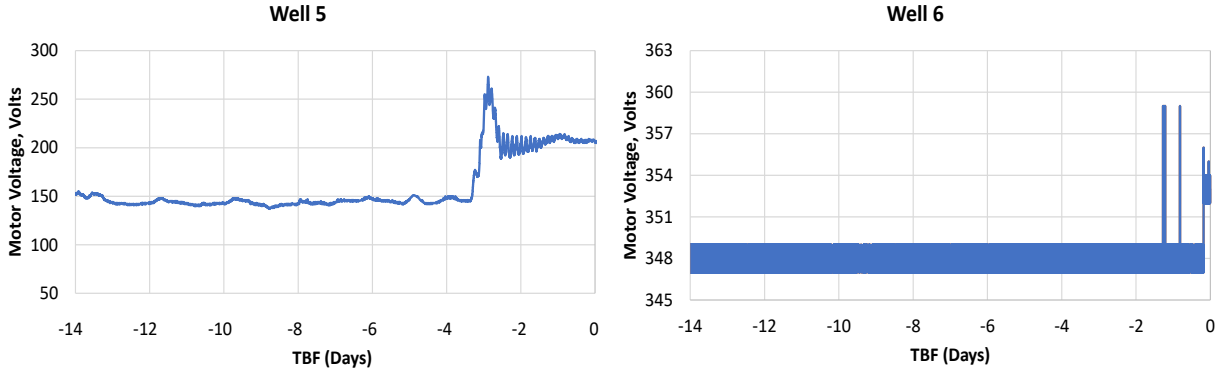

Figure 3-11: Downhole Sensor Failure Specific Failure

v. Broken Pump Shaft:

The pump shaft is long and thin, carrying a significant axial thrust. A slight radial stability change can cause considerable vibrations and ultimately rupture the shaft. The shafts must be robust enough to transfer motor torque to pump stage impellers (Alhanati *et al*., 2001). Pump shaft breakage may be caused by a variety of factors, including pump operation away from its design point, manufacturing faults, unbalanced impeller or rotor, fluid properties (such as high viscosity fluids), pump-driver misalignment, and severe vibration (Takacs, 2009). The typical causes of

severe vibration are cavitation, exceeding the critical speed, passing vane frequency, and operation outside of the optimal efficiency range. As a result, the pump's bearings deteriorate, enabling the shaft to bend and finally break. Due to the absence of vibration parameters, detecting the symptoms proceeding broken shaft failures is not possible based on the available data of this study. Hence, no broken shaft early warning signs were detected from the available data.

3.3.4 SFMs Summary:

Table 3-7 summarizes the described ESP specific failure modes with their early warning signs in the ESP parameters. The average time before the failure (TBF), when each sign starts to appear, is also included based on the ESP failure data of this study. These qualitative observations can be used as a general guideline to identify the ESP operational issues earlier and with less effort.

Table 3-7: Each ESP SFM with their Early Warning Signs, Based on the Data of This Study

SFM	Early Signs	Average TBF (Days)	Standard Deviation of TBF
Broken Shaft	Vibration data not available	-	-
High Motor Temperature	Motor current increases	1.1	0.64
	Motor voltage increase	1.7	1.56
	Motor temperature increases	0.8	1.83
High Cable Temperature	Motor voltage increases	1.1	1.11
	Motor current increase	1	0.92
Overload	Wellhead pressure drops	3.1	1.22
	Motor current increase	2.4	0.53
	Motor temperature increase	2.4	0.94
Plugged with sand	Hard start followed by motor current spikes	2.5	0.66
Power Fail	Voltage spikes	1.85	1.35
	Current increases	2.9	0.91
	Wellhead pressure drops	0.7	1.45
Sensor Failure	Motor Voltage spikes and continue increasing	1.5	1.62
Under Voltage	Motor voltage drops	3.7	0.43
	Motor current Increases	1.4	0.75
Underload	Pump intake pressure increases	2.5	0.65
	Motor current decreases	3.9	0.59

3.4 Case Studies

In this section, two specific case studies will be presented in detail to discuss the failures and their corresponding effects on the ESP and the well's production. The ESP in Well #1 experienced electrical failure, specifically phase unbalance. Well #2, however, experienced electrical failures induced by overcurrent. Both cases share a common general reason for failure but have different underlying causes of failure. Each case includes the failed ESP's history, outcomes of the Dismantle Inspection and Failure Analysis (DIFA), and the underlying cause of failure.

3.4.1 Well #1 Case Study

A downhole ESP pump is being used to produce oil from this vertical oil well with API gravity of 29°, average liquid production of 2,200 BPD, 76% water cut, and a gas-oil ratio (GOR) of 545 SCF/bbl. This ESP was first installed and run in the well for 290 days (9 months). Figure 3-12 depicts the wellbore schematic for Well #1. The system was suffering from unstable intake pressure (P_i) and fluctuations in all running parameters (wellhead pressure, flowline pressure, Vibration, Amps, etc.). A week before the failure, the downhole sensor readings were lost, and the system tripped on overload after around eight hours. The preliminary cause of failure for this ESP was diagnosed as electrical failure resulting in a burn in the Motor Lead Extension (MLE) as seen in Figure 3-13. The rest of the system is still good.

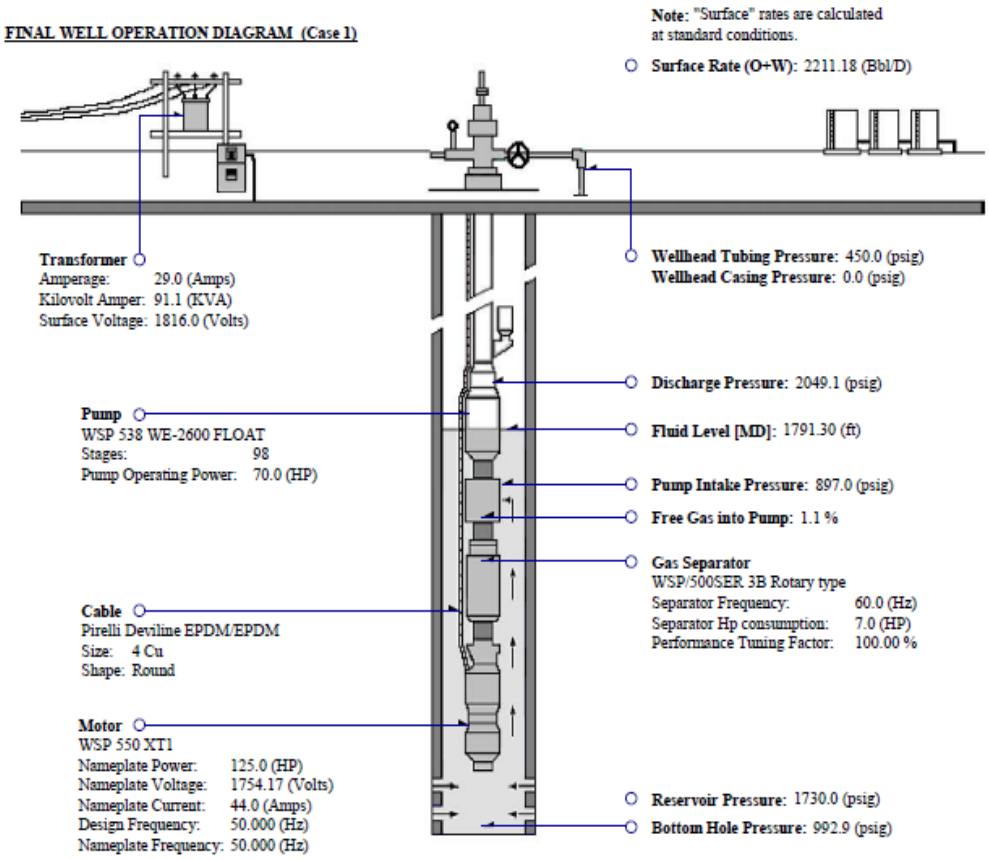

Figure 3-12: Well #1 Wellbore Schematic

During the DIFA, a hole was discovered on the housing at the 8th stage from the top. The housing was dismantled, and a severe wear was discovered on all stages, as well as a scale-like material on the outside of the diffuser, emulsion/foamy fluid on the outside of the top 16 diffusers, and a hole on the diffuser due to severe wear. Figure 3-14 shows the pictures of multiple stages with the above issues highlighted. With regards to electronics, the MLE cable was blown out 24 ft above the pothead, and the main power cable had a thick layer of scale-like material on it, as shown in Figure 3-13. Apart from the MLE cable, the cable was verified to be electronically sound. The rest of the system was also still functional.

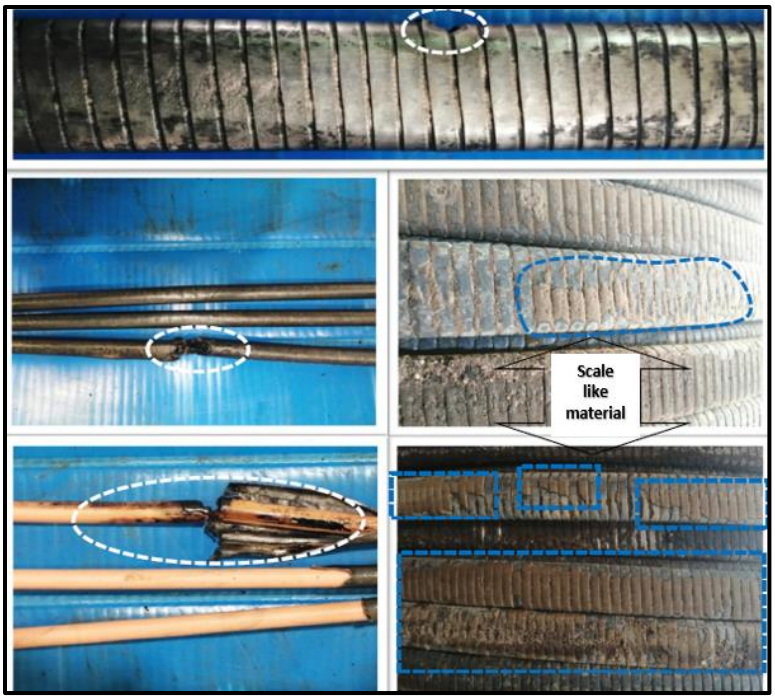

Figure 3-13: MLE Cable Dismantle Findings (Well #1)

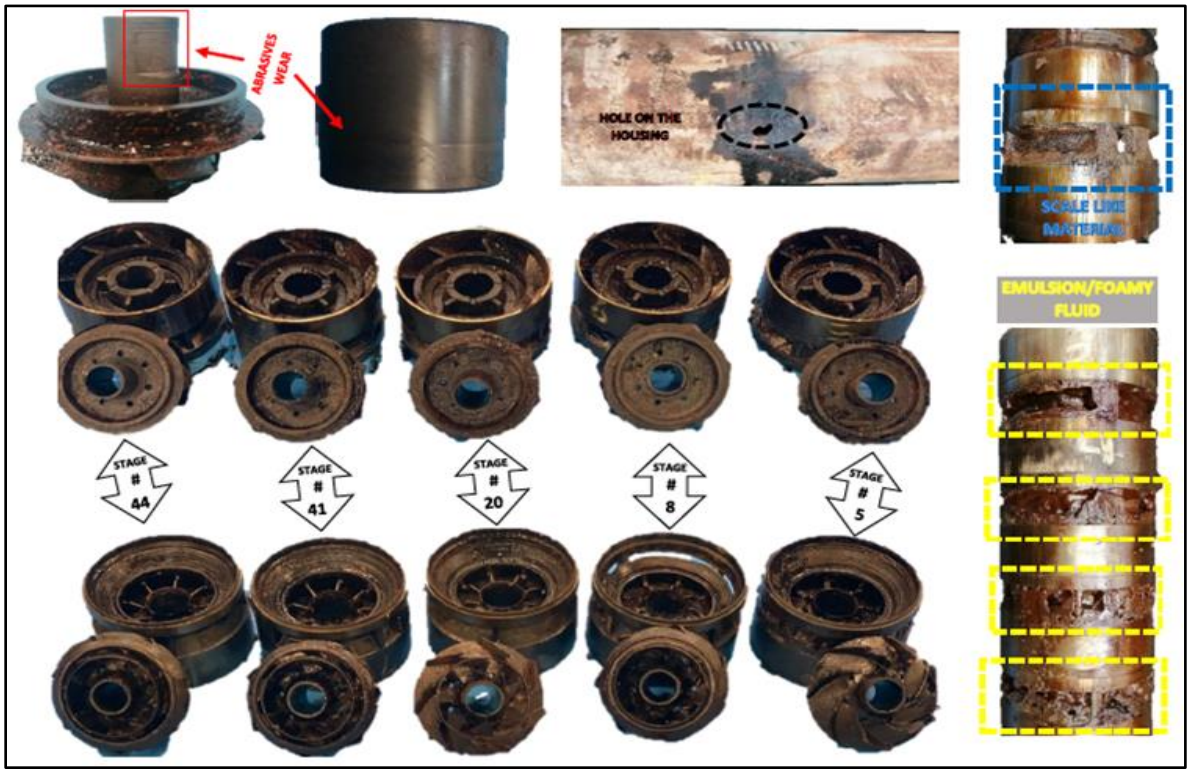

Figure 3-14: Findings from the Dismantled Pump (Well #1)

Many factors contributed to the wear of the pump's stages over its lifespan, including fluid type, scale/solids, backpressure, and down-thrust. The formation's scale-like material caused severe wear, resulting in power overload. Overload increased current draw, resulting in overheating and the cable blowing out. This wear caused a hole in the pump, and as a result, the MLE cable at the hole location was shorted. Table 3-8 summarizes the failure analysis results for Well #1 based on the DIFA report. The root cause of the failure in this situation was found to be the well's Reservoir Performance. A workover was performed to selectively re-perforate the formation based on Pulsed Neutron Capture (PNC) log results for water saturation and hydrocarbon layers. The well was then re-completed with a new ESP.

Table 3-8: Well #1 Failure Analysis Clarifications

Reason for Pull – General:	Electrical
Reason for Pull – Specific:	Phase Unbalance
Primary Failed Item:	ESP Cable
Secondary Failed Item:	Motor Lead Extension
General Failed Descriptor:	Electrical
Specific Failed Descriptor:	Short Circuit
General Failure Cause:	Reservoir or Fluids
Specific Failure Cause:	Reservoir Performance

3.4.2 Well #2 Case Study

A downhole ESP pump is being used to produce oil from a vertical oil well with API gravity of 30°, average liquid production of 3,000 BPD, 60% water cut, and a gas-oil ratio (GOR) of 428 SCF/bbl. This well was run with an ESP for 102 days (3 months). Figure 3-15 depicts the wellbore schematic for Well #2.

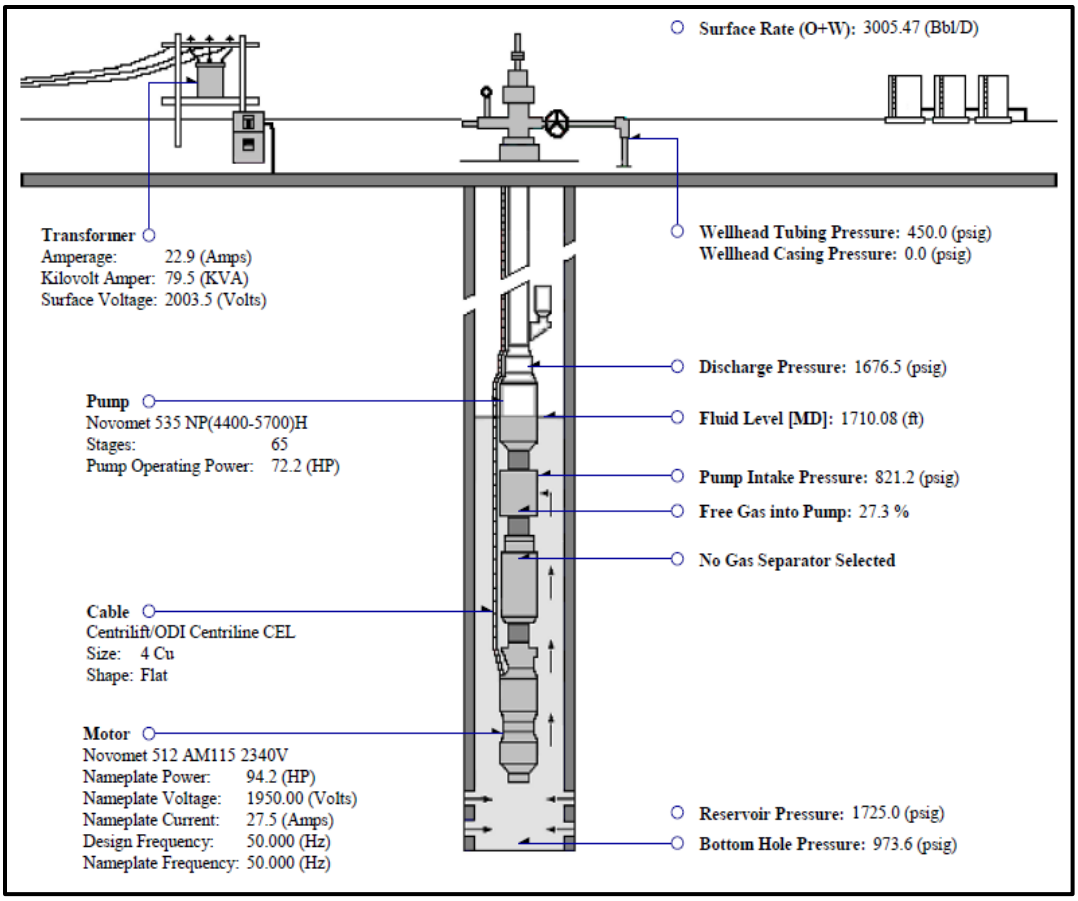

Figure 3-15: Well #2 Wellbore Schematic

The well faced electrical failure due to overcurrent. During the pull out of hole (POOH) process, it was discovered that the pump shaft was stuck, the motor was unbalanced and grounded. In addition, the ESP cable was split and received in two parts. The lower side had satisfactory insulation resistance results, while the upper section suffered considerable mechanical damage during the POOH process, which damaged the cable's insulation layers. The remaining equipment were in good condition.

The motor had two melted leads and was grounded, as shown in Figure 3-16. During the POOH, the cable suffered some mechanical damage, and its upper side had less insulation resistance, as shown in Figure 3-17. The impellers and sleeves of the pump showed radial wear,

as shown in Figure 3-18, with some fallback debris on the upper side. In addition, the pump stage washers had light upthrust wear.

Table 3-9 summarizes the Failure Analysis results for Well #2, based on the DIFA report. The most likely root cause of failure was diagnosed as filter clogging with foreign material, resulting in reduced flow along the items inside the shroud and poor cooling. On the housing, there was an overheat warning, indicating that the Motor had overheated, resulting in electronic failure. As a result, a gravel pack was advised for installation.

Table 3-9: Well #2 Failure Analysis Clarifications

Reason for Pull – General:	Electrical
Reason for Pull – Specific:	Overcurrent
Primary Failed Item:	ESP Motor
Secondary Failed Item:	Motor End Connectors (Y-point/Leads)
General Failed Descriptor:	Material
Specific Failed Descriptor:	Melted / Damaged
General Failure Cause:	Reservoir/Fluids
Specific Failure Cause:	Sand/ Mud

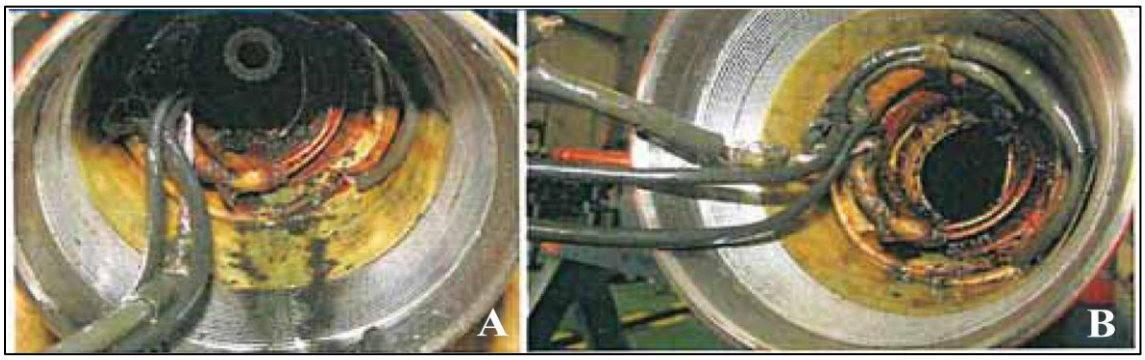

Figure 3-16: Motor Dismantle Observations (Well #2). A) upper side terminal leads melted, B) upper side terminal leads cleaned out.

Figure 3-17: MLE Cable Dismantle Observations (Well #2), A) cable's external condition, B) cable armor's condition, C) a spot in the armor at 935' from MLE, D) opening in the cable armor, the lead insulation pressed.

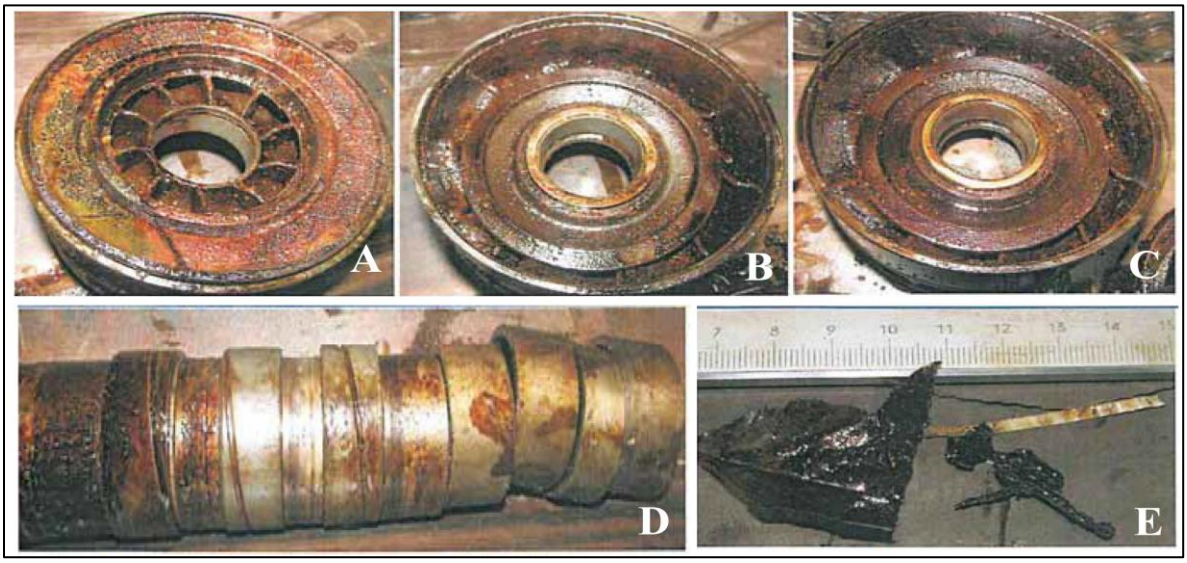

Figure 3-18: Pump Stages Dismantle Findings (Well #2), A) base diffuser's condition, B) middle diffuser's condition, C) head diffuser's wear, D) sleeve's condition at the head, E) debris found in the pump head.

3.5 The need for Machine Learning

Due to a variety of factors such as large gas loads, elevated temperatures, and corrosive conditions, ESP output usually deteriorates steadily, eventually resulting in service interruption (Gupta *et al.*, 2016b). Failure of an ESP has a significant financial impact due to lost production and replacement or intervention expenditures. In 2012, ESPs accounted for 49% of the $10.9 billion artificial lift market, according to Spears & Associates (Donner *et al.*, 2014). In a study,

GE Oil & Gas estimated the annual expense of both lost production and intervention due to ESP shutdowns. Their estimated cost of lost production was up to $3 million, while the cost of intervention could be up to $1 million per well per year (Carrillo, 2013). Their conclusions and findings are summarized Table 3-10.

Table 3-10: Assumptions and Results from a GE Study on ESP failure costs (Carrillo, 2013)

Lost production Costs	Price of oil barrel	$100
	Typical production	500 bbl/d
	Water cut	70%
	Estimated downtime	2 days
	Estimated incidents/year	10
	Estimated costs	500 bbl/d × 20 × 0.3 × $100 = $3MM
Intervention Costs	Onshore conventional wells	$5K to $25K per intervention
	Onshore unconventional wells	$150K to $250K
	Offshore wells	Up to $1MM

Due to such enormous costs for ESP failures, operators are progressively investing in real-time monitoring systems that use the downhole data to track the ESP performance and provide alerts for tripping or failure occurrences. Unfortunately, such systems act only after an incident happens. Hence, there is a necessity to harness the huge amounts of data received in real time from ESP operations to build solutions that shift from a reactive to a proactive approach (Brulé, 2013). With both static and dynamic data being collected in real-time, the energy industry now has data-driven approaches at its disposal to improve the efficiency of its operations (Bravo *et al*., 2014). These solutions aid in early detection of failures, identification of reasons, and recommendation of corrective actions (Gupta *et al*., 2016c).

Operators may avoid failures, reduce downtime, and increase the lifetime of ESP's by implementing analytical models to identify ESP failures. These models may be based on the first principal approaches incorporated into monitoring procedures, thereby supporting, and enhancing decision-making (Stone, 2007). Additional analytics-based Artificial lift monitoring models will

become viable in the future, assisting operators in their goal of shifting from reactive event-based maintenance to proactive predictive maintenance of artificial lift operations (Gupta *et al*., 2016c). Chapters 5 and 6 describe my attempt to reach this objective using the available tools of this study.

Chapter 4 ESP Reliability and Financial Evaluation

This chapter discusses the ESP reliability analysis and the financial consequences of ESP failures. The probability of ESP failure is a crucial indicator in reliability studies. Accurate assessment and forecast of ESP failure probability aids in a better understanding of financial consequences and helps to prioritize a well's intervention requirements.

4.1 ESP Reliability

Mean time between failures (MTBF) quantifies the average time required for a system to fail. MTBF measurements may be used to evaluate the ESP performance, design, and reliability. The MTBF is defined as the arithmetic mean value of the reliability function, *R(t)*, which may be written as the predicted value of the time till failure density function, *f(t)* (Birolini, 2018).

$$MTBF = \int_0^\infty R(t)dt = \int_0^\infty t\, f(t)\, dt \qquad \textbf{4-1}$$

Where $R(t)$ is the reliability function and f(t) is the failure density function

$$f(t) = \lambda e^{-\lambda t} \qquad \textbf{4-2}$$

Where λ is the failure rate

For a given well, Equation 4-3 may be used to calculate the MTBF as the average time between consecutive failures. In addition, Equation 4-4 may be used to calculate the failure rate.

$$MTBF = \frac{\sum(Start\ of\ Downtime\ of\ Current\ Failure - Start\ of\ Uptime\ after\ last\ Failure)}{Number\ of\ failures} \qquad \textbf{4-3}$$

$$Failure\ Rate = \frac{1}{MTBF} \times 100 \qquad \textbf{4-4}$$

The ESP's MTBF trend for the 10 studied wells over the five-year period is shown Figure 4-1. All wells have their maximum MTBF values initially, with times ranging between 46 and 354 days. In 2017, Well 1 has the longest MTBF of 354.5 days, while the lowest was 46 days for Well 6. The MTBF decreases with time, particularly in the third year of operation (2019), until it reaches its lowest value in 2021. In this year, Well 8 has the longest MTBF of 38 days, whereas Well 3 has the shortest duration of 10.93 days.

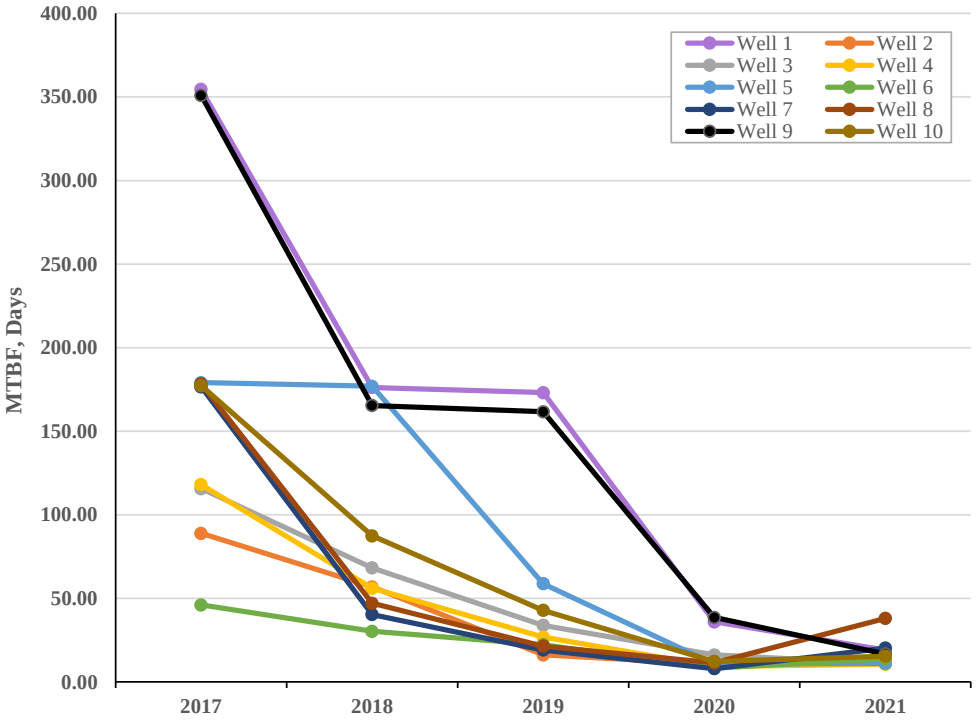

Figure 4-1: ESP's MTBF for All Wells Over 5 Years of Production

The failure rates of ESPs in the 10 wells throughout the five-year production period are shown in Figure 4-2. The trends are inverse of the trends observed in the previous figure for MTBF. The failure rate increases each year until it reaches a peak of 12.6% in the fourth year (2020). The failure rates of Wells 6 and Well 7 are consistently greater than those of the other wells.

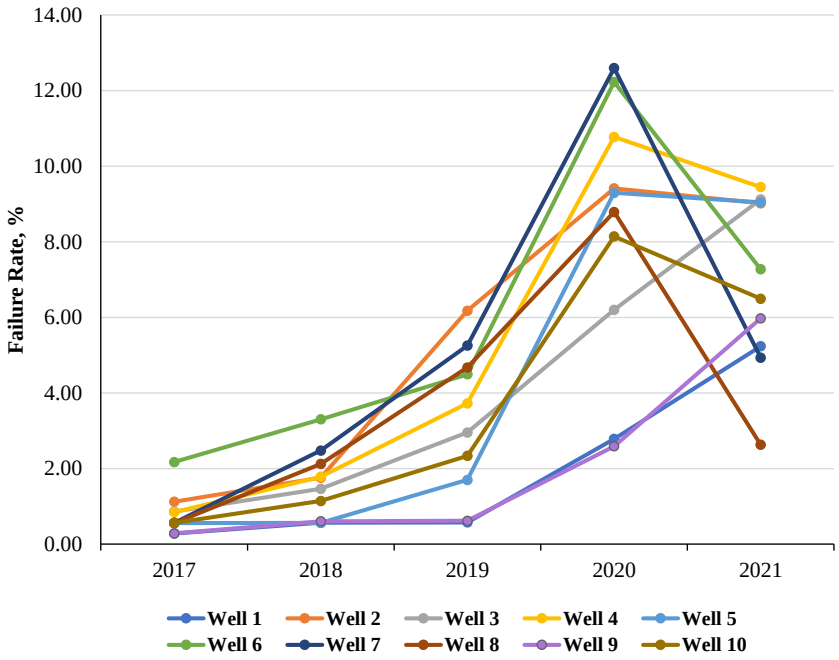

Figure 4-2: ESP Failure Rates with Time for the 10 Tested Wells

In this chapter, I am focusing on wells 6 and 7 due to their higher failure rates. Figure 4-3 depicts the general failure categories for Well 6, with electrical failures accounting for 80% of the failures, 15% for reservoir, and 5% for gas effects. Figure 4-4 shows the specific failure modes by subcategory. Power failure accounts for 71% of all SFMs, while the remaining 29% is almost evenly split across MTR cable, overload, undervoltage, and underload.

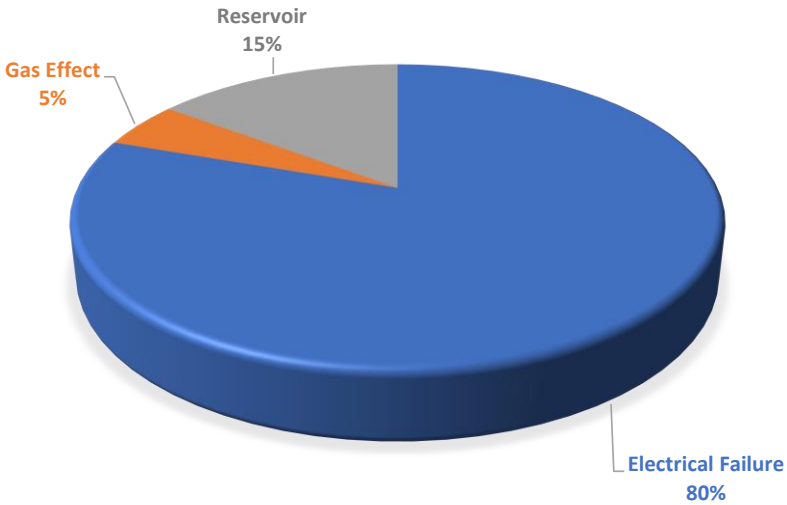

Figure 4-3: Well 6 – ESP's General Failure Modes Over the 5 Years of Production

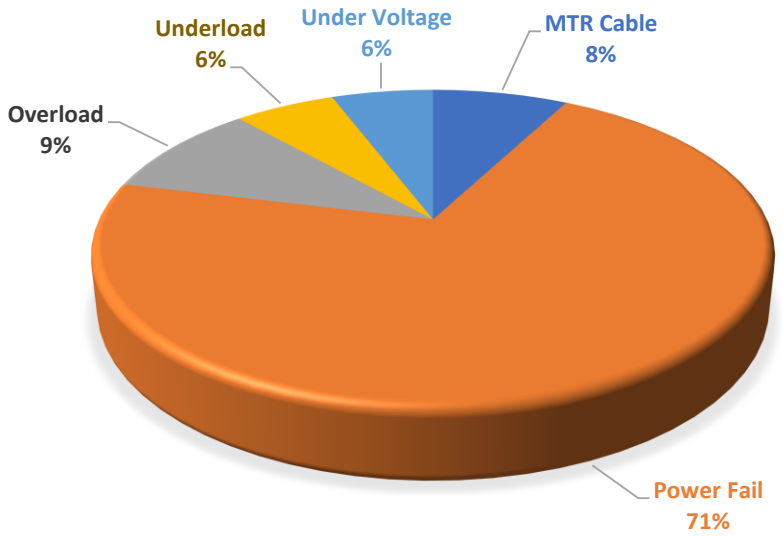

Figure 4-4: Well 6 – ESP's Specific Failure Modes (SFM's) Over the 5 Years of Production

Figure 4-5 depicts the general failure categories for Well 7, with electrical failures accounting for 86% of the failures. Figure 4-6 shows the specific failure modes by subcategory, with power failure accounting for 72% of all SFMs, while overload is 15%, undervoltage is 9%, and overvoltage is 4%.

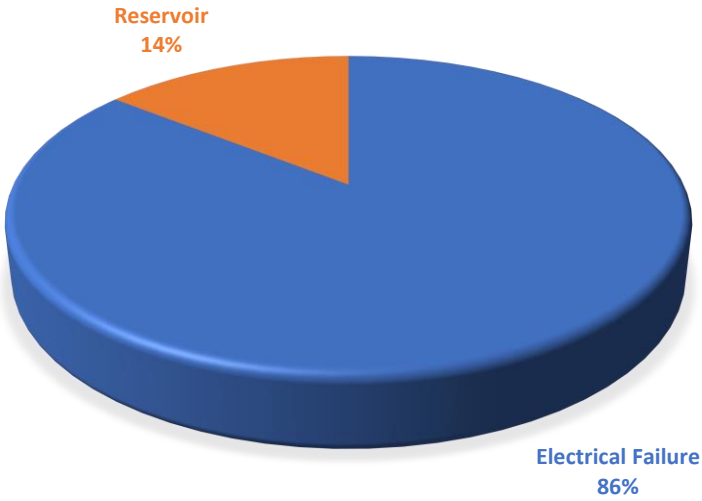

Figure 4-5: Well 7 – ESP's General Failures Over the 5 Years of Production

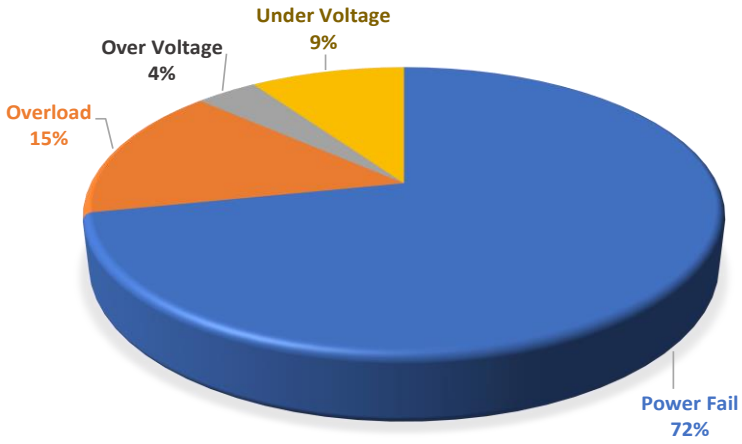

Figure 4-6: Well 7 – ESP's Specific Failure Modes (SFM) Over the 5 Years of Production

4.2 Weibull Analysis

Failure data may be modelled using a wide variety of distributions, including normal, exponential, Rayleigh, Weibull, gamma, lognormal, and along with others (Sawaryn *et al.*, 2002). Weibull Analysis is an efficient statistical data analysis technique for assessing the reliability features and trends of failure data. The two-parameter Weibull distribution is defined as follows (Sawaryn *et al.*, 2002):

$$f(t) = \frac{\beta}{\eta}\left(\frac{t}{\eta}\right)^{\beta-1} e^{-\left(x/\eta\right)^{\beta}} \qquad \textbf{4-5}$$

Where β is the shape factor, η is the scale parameter, and t represents the value to be evaluated. The shape factor (β) determines the distribution's behaviour. The failure rate drops with time if $\beta < 1$, meaning that the ESP gets more trustworthy as it matures. Manufacturing or installation errors may be a cause of this. The failure rate increases if $\beta > 1$, which is commonly due to pump wear. Finally, if β =1, it denotes a consistent failure rate that is time independent.

Based on the field data from the 10 wells, a Weibull model was built to estimate the probability of failure. The previously estimated MTBF values were utilized as inputs to the Weibull analysis model. The Weibull shape and scale parameters were estimated using Median

Rank Regression. Then the Weibull Probability plots were generated. These plots depict unreliability on a logarithmic scale, which is defined as the probability of failure (%) vs. time after the previous failure (days). The legend indicates the number of MTBF points (5 points for five years), Weibull parameter estimates (β and η), confidence limits (Beta-Binomial Bounds), and the degree of confidence (90%).

Figure 4-7 and Figure 4-8 illustrate Well 6's and Well 7's unreliability plots, respectively. The Weibull line accurately fits the MTBF points in both plots with R^2 values of 99.4% and 86.5%, respectively, and low p-values. The predicted Weibull parameters for Well 6 are β of 1.538 and η of 28. The $\beta > 1$ value is a sign of pump wear and increasing failure rate with time. According to the plot, the ESP in Well 6 has a 50% probability of failure 20 days after the last failure, 75% probability of failure 35 days after it, and 90% probability of failure 75 days afterwards. The maximum probability of failure occurs 80 days after the previous failure with 99% probability.

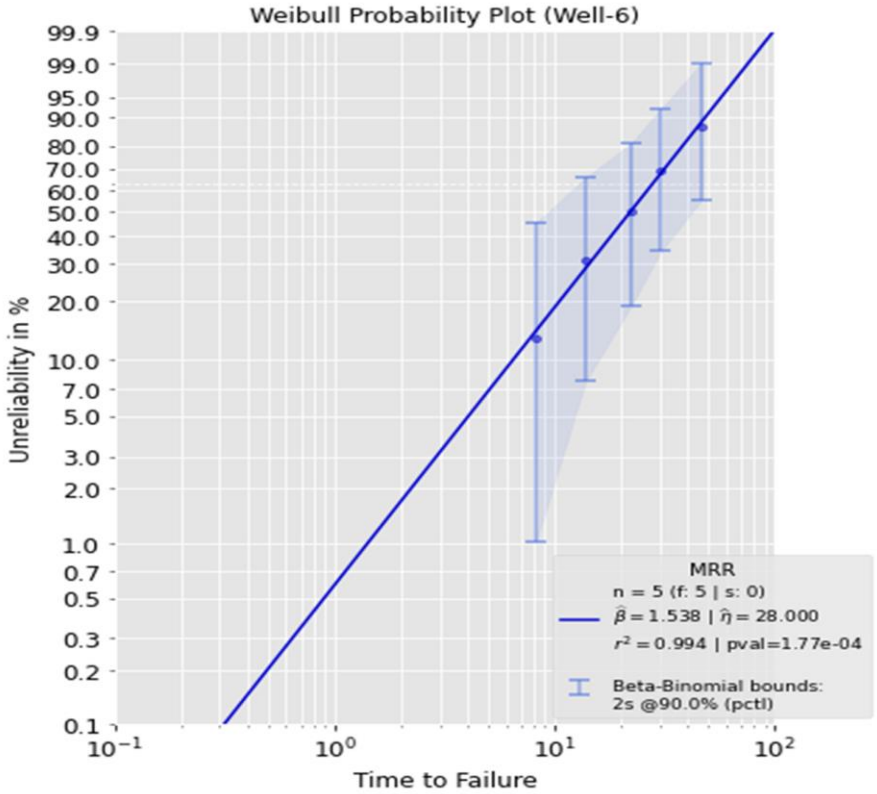

Figure 4-7: Weibull ProbabilityPlot of Well 6

From Figure 4-8, the predicted Weibull parameters for Well 7 are β of 0.835 and η of 52.87, a possible indicator of manufacturer errors. According to the plot, the ESP has a 50% probability of failure 32 days after the last failure, 75% probability of failure 65 days afterwards, and 90% probability of failure after 190 days. The maximum predicted failure probability occurs in 300 days with 99% chance. Comparing both ESP wells, well 7 is more reliable than well 6, as the 90 percent probability of failure occurs after 190 days vs. 75 days in well 6. The Weibull unreliability plots for the remaining wells are shown in Figure A-1 through Figure A-8 in appendix A.

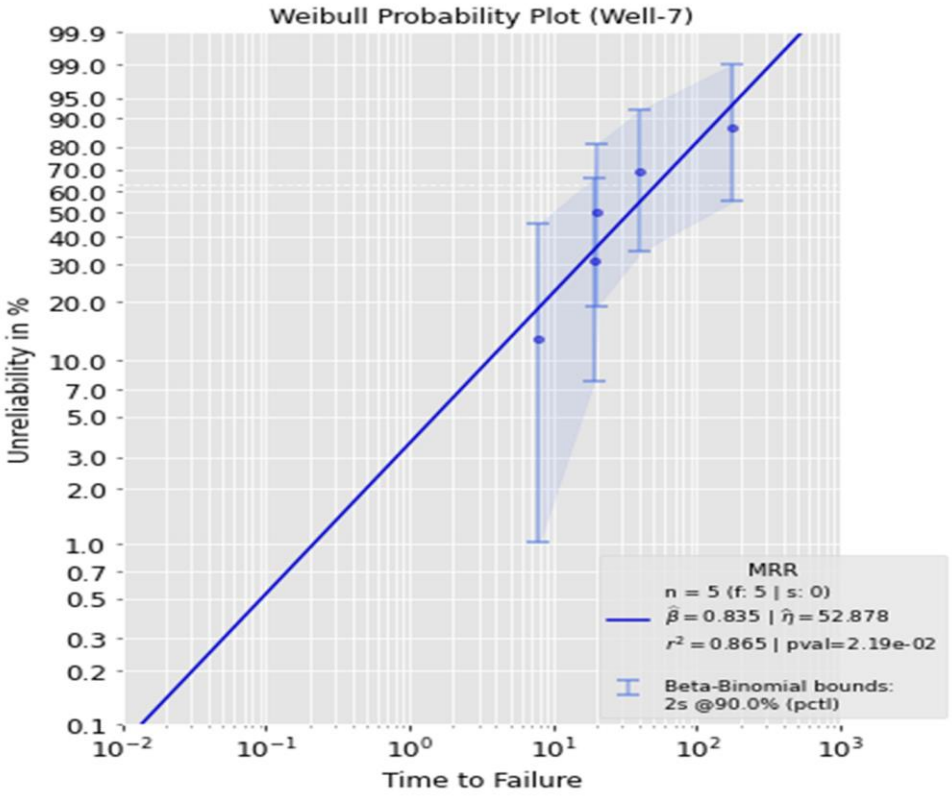

Figure 4-8: Weibull Probability Plot of Well 7

4.3 Financial Evaluation of Well Interventions

There are very few wells that can produce continuously from the time they are drilled until the time they are shut permanently. The system begins to fail when the pump, seals, tubulars,

sensors, and formation pressures deteriorate. Operators depend on well intervention to handle these issues. There are two general types of interventions, light and heavy. Workers descend equipment or sensors into a live well while maintaining surface pressure for minor operations. On the other hand, there are times when heavy interventions for large modifications in the well structure result in the well's suspension of production.

Engineers must evaluate the expense of intervention against the benefit of greater oil production when deciding whether to intervene. Typically, intervention is prioritized for the well with the greatest oil production rate, providing additional revenue for the whole oil field. The field engineer usually creates a list to prioritize the intervention timetable for the down wells. These lists are constructed depending on the well's production rate, downtime duration, and needed well testing (if any) while waiting for the intervention.

This section provides financial assessment methods for assessing the financial effect of ESP failure interventions and improvement decision making. Net cash flow, discounted cash flow, and net present value are the most common metrics to achieve these objectives.

1. Net Cash Flow (NCF):

The term "net cash flow" refers to the money gained after all costs have been met for a certain year or period. Annual NCF is calculated by subtracting cash outflows (costs) from cash inflow (revenue) over the course of a year (Equation 4-6).

$$NCF = Revenue\ (Cash\ Inflow) - Cost\ (Cash\ Outflow) \quad \textbf{4-6}$$

Revenue is generated by the sale of crude oil, natural gas, or condensate, and other operations. Since this study is based on Kuwaiti assets, prices for the "Kuwait Export Blend (KEB)" over the past five years are used. The variation in the international market price of KEB is shown in Figure 4-9.

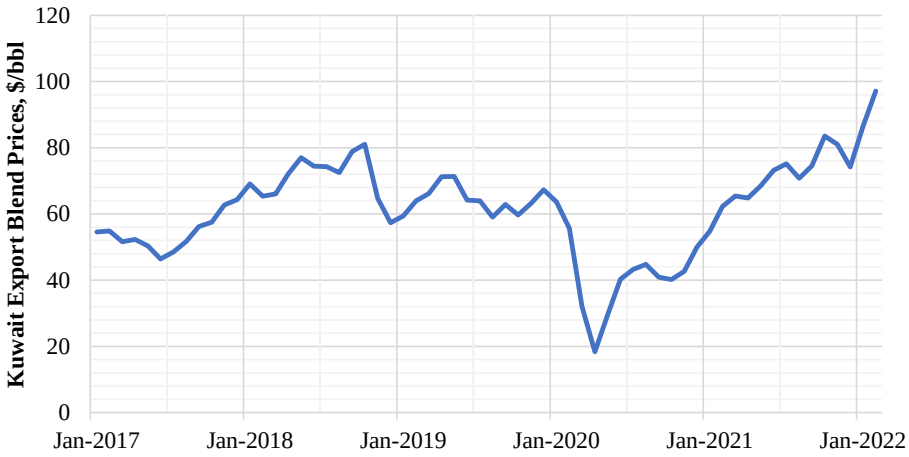

Figure 4-9: Kuwait Export Blend Prices ($/bbl)

There are two types of costs: capital expenditure (CAPEX) and operating expenditures (OPEX). CAPEX refers to exploration and development expenses, whilst OPEX refers to the operational costs. The distribution of field expenditures, CAPEX and OPEX, is influenced by a company's strategy, project objectives, and reservoir type. Because the data of this study are from shallow vertical wells, the anticipated CAPEX and OPEX for each well are estimated as $4 million and $7 per barrel, respectively.

2. Discounted Cash Flow (DCF):

Discounted cash flow is a strategy for converting the time value of money to a present value reference by discounting net cash flows (Mian, 2011). The discounted cash flow at the end of the year is denoted by the following:

$$DCF_y = \frac{NCF_y}{(1+i)^y} \quad \textbf{4-7}$$

Where NCF_y is the Net Cash flow at the end of year y. i is the discounted rate, while y is the number of years ($y=0,1,2,\dots Y$)

3. Net Present Value (NPV):

The net present value (NPV) of a project is the summation of its discounted yearly cash flows (Mian, 2011). The NPV is expressed as follows:

$$NPV = \sum_{y=0}^{Y} DCF_y \qquad \textbf{4-8}$$

Where DCF_y is the Discounted Cash flow at the end of year *y*. *y* is the number of years (*y=0,1,2,…*)

When the NPV is positive, there will be a return on the project. If the net present value of the project is negative, it indicates that cash inflow will be less than cash outflow. If the NPV equals zero, then the revenues and expenses are identical. The degree of risk involved with a project should be considered before making a decision. As a matter of corporate policy, 10% is the discount rate utilized to compute net present value in financial analysis.

4. Lost Production and Intervention Costs:

As explained in Section 3.5 and using the assumptions in Carrillo's (2013) GE Study on ESP failure costs (Table 3-10), the costs of lost production and intervention are computed. The cost of a workover is estimated to be $66,000 per day for a shallow vertical well. The cost of onshore intervention ranges between $5,000 and $25,000. Carrillo (2013) provided the following estimate for the cost of lost production:

$$Lost\ Production\ Cost\ (\$) = Oil\ Production\ Rate \times \sum Downtime\ \times Oil\ Price \qquad \textbf{4-9}$$

Where Downtime is the duration of shutdown due to failure per year (days), Oil Production is in BPD, and Oil Price is in $/bbl

As shown earlier, Wells 6 and 7 have higher failure rates than the other wells. Their data are utilized in financial analysis to assess the intervention priority for these two wells with highest failure probabilities, distinct oil production rates and comparable mean times between failures. As seen in Figure 4-10, Well 7 produces more oil than Well 6, with around 3,500 bpd on average, compared to 950 bpd for Well 6.

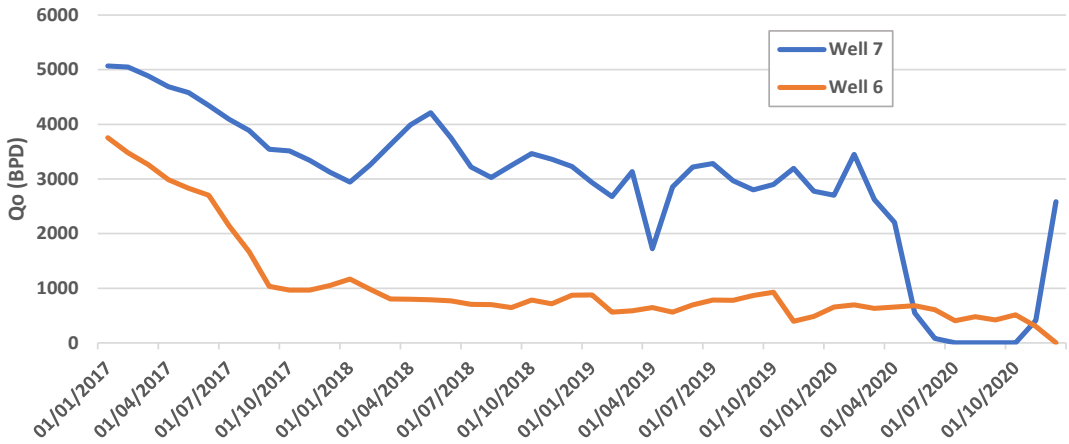

Figure 4-10: A Comparison of the Oil Production Rates of Wells 6 and 7

The cash flow diagrams of the two wells were created using Equation 4-9, 4-6, and 4-7. The cash flow of Well 6 is shown in Figure 4-11, where revenue declines year after year due to an increase in ESP failures in 2019 (3rd year of operation), leading to a shorter MTBF as explained in Section 4.1. As the MTBF decreases, the frequency of ESP failures increases, resulting in an increase in the cost of lost production and intervention costs. As a result, the NCF reduces over time because of increased expenses, as well as a decline in revenue. This results in an overall net present value (NPV) of $45.273 MM for Well 6.

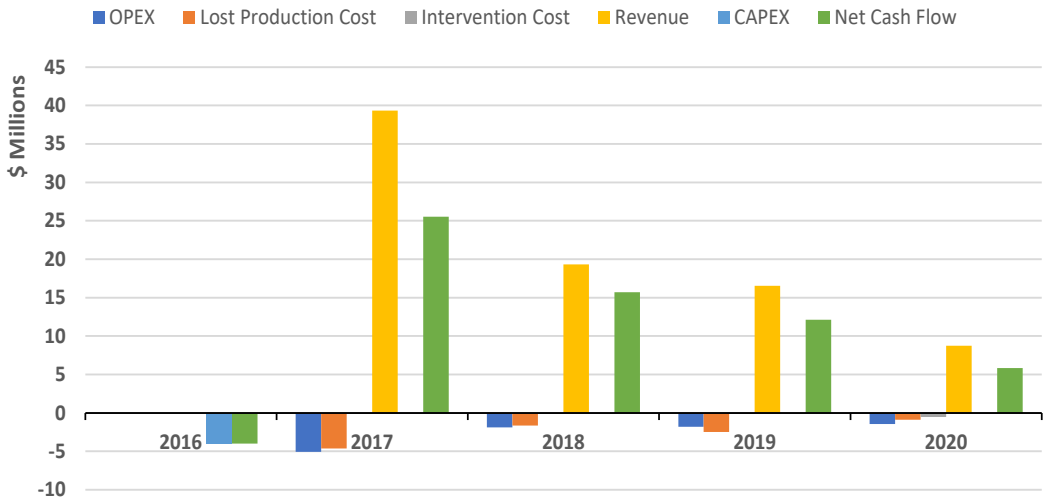

Figure 4-11: Well 6 Cash Flow Diagram

The cash flow for Well 7 is shown in Figure 4-12, where a similar pattern of revenue declines exists owing to the reduced production and an increase in ESP failures (shorter MTBF over time). Because of higher expenditures and a loss in revenue, the NCF decreases over time. This generates a total NPV of $143.210 MM for Well 7.

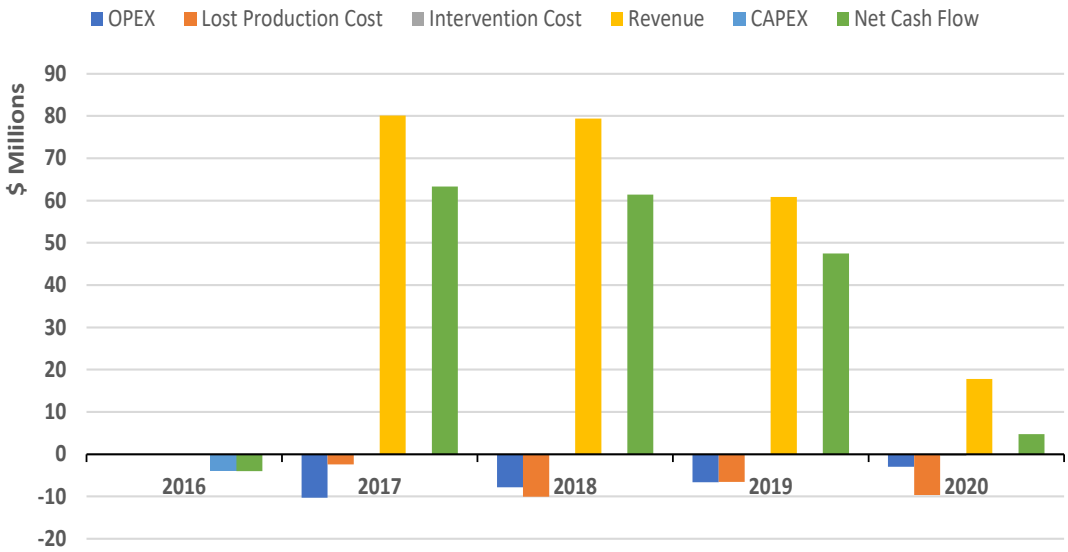

Figure 4-12: Well 7 Cash Flow Diagram

A Multi-Criteria Decision Analysis (MCDA) framework was used as a decision-support tool to assist in ranking wells for intervention. This technique acts as a prioritizing strategy, giving a systematic framework to objectively analyze issues and support the field engineer in making decisions. This strategy is very advantageous when used across an entire field including hundreds of wells that need intervention at the same time, with a limited number of intervention units/ workover rigs.

The Technique for Order Preference by Similarity to Ideal Solution (TOPSIS), a subset of MCDA, was applied. This method selects the alternative with the lowest Euclidean distance to the ideal solution and the largest distance to the negative ideal solution. It is based on the relative weights and influences (positive or negative) of the specified factors.

The TOPSIS tool was applied to the ten oil wells on the assumptions that all wells need intervention concurrently and no well testing is necessary. Using the findings of the Weibull failure probability plots and the financial impact of well intervention, Figure 4-13 was constructed. A comparison of the net present values (NPV), times of 90% unreliability, and anticipated oil productions of all wells is shown in Figure 4-13.

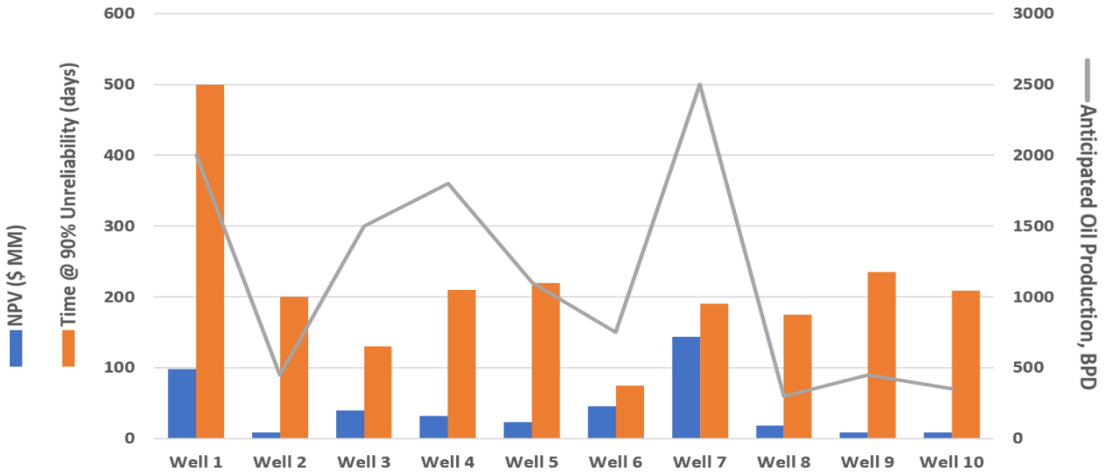

Figure 4-13: Prioritization Factors for Well Interventions

A multiple linear regression model (Least Squares method) was run to determine the weight of each of the introduced factors. Multiple linear regression investigates the linear connection between the dependent variable and independent variables. The dependent variable is NPV, and the independent variables are unreliability time and anticipated oil production. Coefficients are produced with high R^2 and adjusted R^2 values of 86% and 82%, respectively, and a low p-value, as summarized in Table 4-1.

Table 4-1: Multiple Linear Regression Results

Independent Variables	Coefficient	Standard error	t	P>|t|	[0.025	0.975]
Time @ 90% Unreliability (days)	-0.0276	0.059	-0.467	0.653	-0.164	0.109
Anticipated Oil Production, BPD	0.0452	0.01	4.304	0.003	0.021	0.069

The anticipated oil production is proportional to the NPV, due to the significant influence of future revenue on NPV growth. The decision criteria are as follows: wells with a greater net present value, a higher anticipated oil production rate, and a shorter period for 90% unreliability are prioritized. Using the TOPSIS ranking approach, wells are ranked according to the specified decision criteria. The proposed intervention sequence for the wells of this study is depicted in Figure 4-14.

Figure 4-14: Wells Intervention Sequence

Based on the well intervention ranking tool, the field engineer may decide to prioritize the well with the greatest net cash flow for future intervention before others. The recommended intervention sequence is dictated by the Well's NPV and unreliability. To further customize the well intervention ranking tool for a specific field, additional information is required, such as the availability of intervention units/workover rigs on the field, possible need for well testing, the well's location (constructed or unconstructed roads), the weather conditions (such as sandstorms), and the wellsite's readiness for intervention unit/rig (surface lines are disconnected).

Chapter 5 Data Analytics Design

Chapter 5 presents a set of machine learning models (ML) to handle the high frequency (1-minute frequency) of ESP operational data and transform them into actionable information to predict failures. These models allow petroleum engineers to detect problems early, diagnose potential causes, and propose preventative actions. The models are based on field data gathered from the surface and downhole ESP monitoring equipment over five years of production of 10 wells.

Developing and implementing a machine learning model involves several stages, as shown in Figure 5-1. Problem definition, raw data acquisition, data pre-processing and preparation, the construction of various ML models, and introducing the performance metrics are among the ML stages covered in this chapter. In Chapter 6, the evaluation results for each constructed ML model are presented and thoroughly analyzed.

Figure 5-1: Machine Learning Model Workflow

5.1 Data Exploratory Analysis

A preliminary analysis of the ESP operational data shows several interesting trends. First, graphing each variable against time to visualize the data and failure trends is necessary. The wellhead pressure in Figure 5-2 exhibits standard patterns over one year of an oil well's flow, with noticeable spikes owing to ESP pump failure. The analysis is focused on these spikes since they

reflect the ESP failure mechanisms that must be recognized, studied, and forecasted. This pattern is reproduced for all the other ten variables across all the wells.

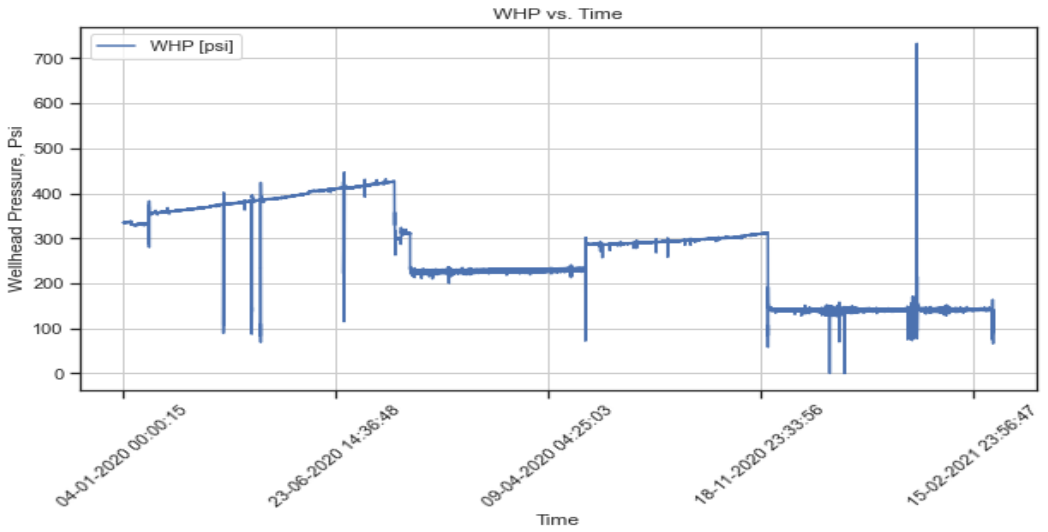

Figure 5-2: Wellhead Pressure with Time (Well #1)

Figure 4-3 provides a statistical summary of the analysed data in this study using boxplots. The plots show a few extreme values that may indicate pump failures and need to be included in the data evaluation. As a result, the dataset includes a substantial number of outliers that have not been eliminated to capture all the data trends during ESP failures.

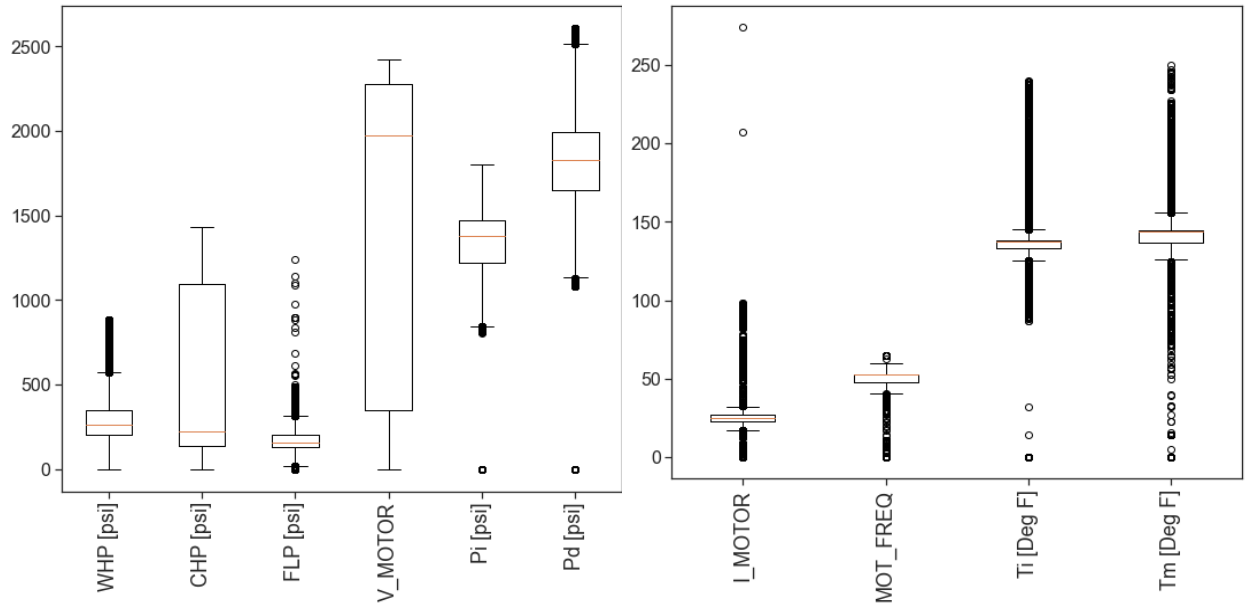

Figure 5-3: Boxplots of ESP Data Variables

A matrix scatter plot showing the relationships between a few of the measured quantities is shown in Figure 5-4. While the relationships are not necessarily linear, there is some association between the variables. For a qualitative analysis of the relationship between the ESP variables, a correlation matrix was created, as shown in Figure 5-5.

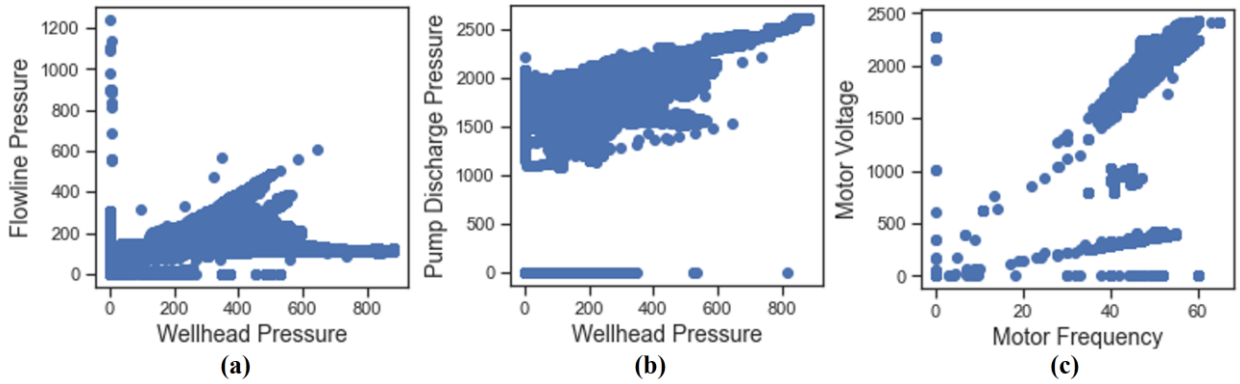

Figure 5-4: Scatter Plots of ESP Data Variables. a) Flowline Pressure vs. Wellhead Pressure, b) Pump Discharge Pressure vs. Wellhead Pressure, and c) Motor Voltage vs. Motor Frequency.

Figure 5-5 depicts a significant positive correlation between the wellhead pressure and the discharge pump pressure, flowline pressure, motor temperature, and motor frequency. The flowline pressure, on the other hand, shows a strong positive correlation with the motor frequency, pump's intake and discharge pressures, suction temperature, and motor temperature. There is also a strong positive correlation between the motor voltage and motor frequency. Motor frequency is shown to be strongly related to the pump's suction and discharge pressures and temperatures. Overall, there are no variables with zero correlation (all positive values), and all the parameters are correlated to some degree.

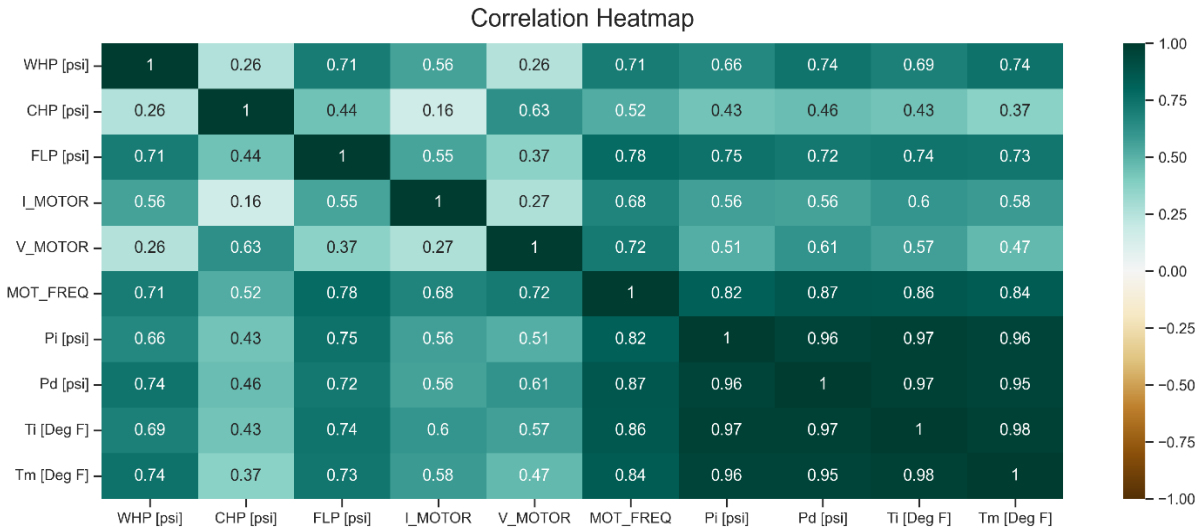

Figure 5-5: ESP Variables Correlation Heatmap Matrix

5.2 Data Pre-processing

The initial stage in developing a machine learning model is data preparation, which serves as a kick-off. High-frequency data are commonly imprecise, inconsistent, erroneous, and devoid of specific attribute values. As a result, data preparation is introduced, assisting in cleaning, organizing, and classifying raw data before usage by the machine learning models.

Missing values are a severe concern in any dataset and can considerably impact the model's quality. Frequently, they are NaNs, blanks, or other notations. Many missing values were identified and eliminated in this research, while ensuring no valuable information was eliminated. Most of the removed data occurred when the variable speed drive (VSD) stopped capturing high-frequency data due to operational concerns. The second most significant issue that arises during data analysis is outliers. An outlier is an observation that deviates from the overall trend. However, the outliers in this study represent the ESP failure modes that must be kept, identified, studied, and forecasted.

The train-test splitting technique is used in this research to evaluate the performances of machine learning algorithms when applied to data not used to train the model. This technique assists in comparing the effectiveness of machine learning algorithms when faced with a predictive

modelling task. At this study, the data was stratified and divided into training and testing datasets, with 80 % of the data used to train the models and the remaining 20 % used to test them.

Machine learning methods perform better when numerical input variables are scaled to a normal distribution. Normalization is a data organization technique used to organize the data in a database without changing the original data distribution. It minimizes duplication and evenly weighs each variable, guaranteeing that no variable dominates model performance because of greater values. It separately adjusts each input variable to the 0-1 range, while preserving value range variations. Hence, data normalization was performed on the entire dataset using the MinMax approach.

Using hyperparameters enhances the performance of machine learning (ML) models; hence, GridSearchCV was also used to optimize hyperparameters. Grid Search-Cross Validation is a strategy that attempts all conceivable parameter value combinations (Grid Search) and delivers the optimum hyperparameters to improve the model performance. When training a model, GridSearchCV not only employs Grid Search, but it also does the cross-validation (K-Fold cross-validation). The K-Fold cross-validation shuffles the data randomly, divides the training dataset into K folds (5 folds), and then trains the model K times, each time excluding a different fold from the training data and utilizing it as a validation set. Consequently, the tuned parameters are used to construct the ML model. The ML model was then evaluated using the test dataset to assess the performances of the ML algorithms.

5.3 Features Generation

Generating physical features is essential for the machine learning model to develop a hybrid failure prediction approach and account for the physical aspects of ESP operations. This

study investigates the gas interference in the ESP system and evaluates pump performance based on the field data. After highlighting the impact of free gas on ESP performance (Section 2.4), this section describes the methods used to estimate the gas volumetric fraction at the intake conditions and details the pump efficiency calculations. A detailed technique is described for estimating the fluid properties based on the field data and using it to calculate the gas void fraction in the pump.

For the ESP efficiency estimation, the general form of pump efficiency (Takacs, 2017) is defined as:

$$\eta_{pump} = \frac{Hydraulic\ Horse\ Power\ (HHP)}{Total\ Input\ Electrical\ Power} \times 100 \quad \textbf{5-1}$$

$$HHP = 7.37 \times 10^{-6} Q_o H_{req} \gamma_o \quad \textbf{5-2}$$

$$Total\ Input\ Electrical\ Power\ = Motor\ Current \times Motor\ Voltage \times 1.341 \quad \textbf{5-3}$$

Where *HHP* and total input electrical power are expressed in horsepower (HP), oil flow rate (Q_o) is expressed in BPD, and 1.341 is a conversion factor to convert kW to HP. Additionally, the motor current (I_m) is expressed in Amperes (A) and the motor voltage (V_m) is expressed in Volts (V). H_{req} is defined as follows:

$$H_{req} = \frac{\Delta P_p}{0.433\,\gamma_o} \quad \textbf{5-4}$$

By rearranging Equation 5-4 and substituting Equation 5-2, 5-3, and 5-4 into Equation 5-1, the efficiency of the ESP pump becomes:

$$\eta_{pump} = \frac{7.37 \times 10^{-6} Q_o \Delta P_p}{0.433\ \times 1.341 \times\ I_m V_m} \times 100 \quad \textbf{5-5}$$

The daily production data were utilized to calculate the efficiency of the ESP pump using Equation 5-5.

As discussed in Chapter 2 (Section 2.4), when evaluating pump performance in multiphase flow, the Turpin correlation is a commonly utilized correlation. This correlation was used to analyze the ESP data in this research. Estimating the Turpin correlation (ϕ) parameter (Equation 5-9) requires the gas volumetric flow rate at the pump intake (q_s), liquid volumetric flow rate at the pump intake (Q), and the gas volumetric fraction (λ_c). Under the pump intake conditions, these parameters can be estimated as follows:

$$q_s = \frac{Q_o\,(GOR - R_s)B_g}{5.615} \quad \textbf{5-6}$$

Where q_s is the free gas rate in RB/D, Q_o in STB/D, GOR and R_s in SCF/STB, and B_g in ft^3/SCF.

$$Q = Q_o\,B_o \quad \textbf{5-7}$$

Where Q is the oil rate in RB/D and B_o in RB/STB

$$GVF = \frac{q_s}{q_s + Q} \times 100 = \lambda_c \quad \textbf{5-8}$$

$$\Phi = \frac{2000\,q_s}{3\,P_i Q} \quad \textbf{5-9}$$

Where Φ is in psi^{-1}

$\Phi < 1$ indicates stable operating region, and $\Phi > 1$ indicates unstable operating region (severe gas interference)

In order to evaluate gas interference in the pump, a set of fluid properties are required. These properties are determined based on the field data, as described below.

5.3.1 Fluid Characteristics Estimation

The ESP data were obtained from wells classified as black oil reservoirs (Figure 5-6). Hence, Black oil models were used to estimate in-situ fluid properties. The range of physical parameters utilized as inputs is shown in Table 3-2. These parameters are used to calculate the fluid properties at pressures above and below the bubble point. Each of the fluid characteristics listed below is assessed at ESP's intake pressure, P_i, and temperature, T_i.

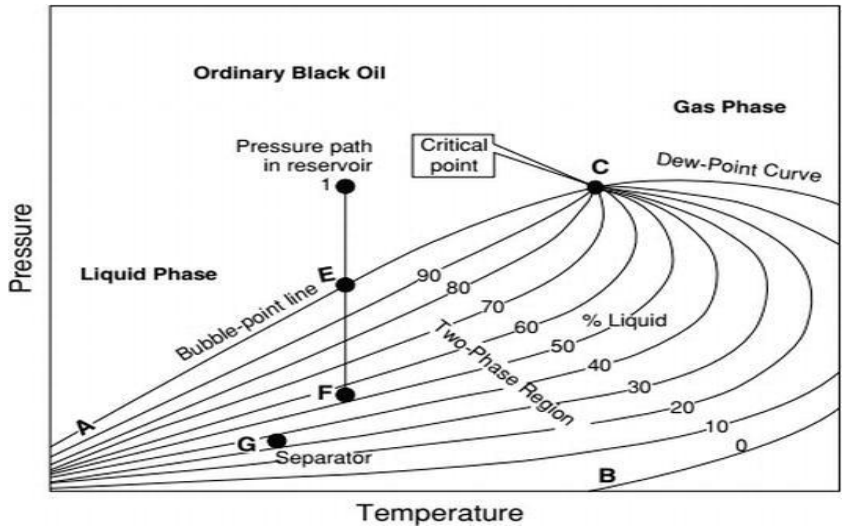

Figure 5-6: Typical Black Oil Model Phase Diagram (Ahmed, 2010)

Several correlations are used to compute the oil formation volume factor (B_o) in a black oil reservoir. Standing (1947) developed one of the most widely used correlations for determining B_o using the fluid temperature, solution gas-oil ratio at bubble point pressure, and oil and gas gravities. This method yields accurate results at pressures above the bubble point. Numerous following researchers, most notably Velarde (1996), attempted to enhance the Standing correlation by considering pressures below the bubble point ($P < P_b$). Standing (1981) and Velarde (1996) both assert that the oil formation volume factor (B_o) is strongly related to the solution gas-oil-ratio (R_s), reservoir temperature (T_{res}), gas specific gravity (γ_g), and oil API gravity. The following fluid properties were derived using the field data listed in Table 3-1.

i. Determination of Solution Gas Oil Ratio (R_s):

The solution gas-oil ratio (R_s) or gas solubility refers to the quantity of gas dissolved in oil. It ranges between 0 to nearly 2,000 SCF/ STB in black oil systems. At a given reservoir temperature, solution gas-oil ratio stays constant at pressures greater than the bubble point pressure ($P > P_b$), indicating that all accessible gases are dissolved in the oil and the gas solubility is at its

maximum value ($R_s = R_{si}$). When the pressure drops below P_b, the gas starts to evolve from the solution, increasing its saturation and decreasing the R_S value, as illustrated in Figure 5-7.

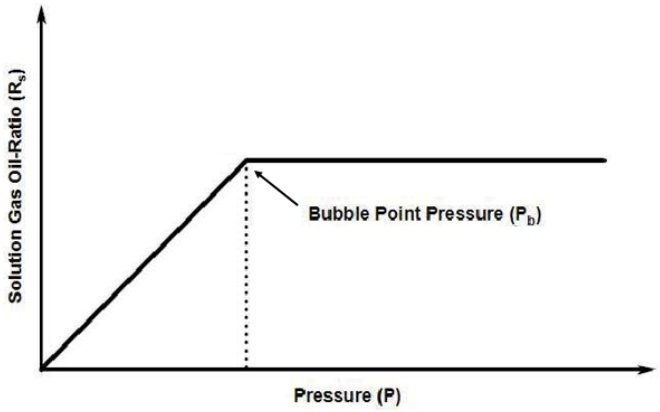

Figure 5-7: Typical Solution Gas-Oil-Ratio Diagram (Ahmed, 2006)

The solution gas-oil ratio below P_b is estimated using the Velarde (1996) correlation. First, the reduced pressure (P_r) and the reduced solution gas-oil-ratio (R_{sr}) are calculated using Equation 5-10 and Equation 5-11, respectively.

$$P_r = \frac{P_i}{P_b} \quad \textbf{5-10}$$

$$R_{sr} = a_1 {P_r}^{a_2} + (1 - a_1){P_r}^{a_3} \quad \textbf{5-11}$$

Where P_r is dimensionless, R_{sr} is in SCF/STB, and $a_1, a_2,$ and a_3 are as follows:

$$a_1 = 9.73 \times 10^{-7}\, {\gamma_g}^{1.672608}\, API^{0.929870} {T_i}^{0.247235} {P_b}^{1.056052} \quad \textbf{5-12}$$

$$a_2 = 0.022339\, {\gamma_g}^{-1.00475}\, API^{0.337711} {T_i}^{0.132795} {P_b}^{0.302065} \quad \textbf{5-13}$$

$$a_3 = 0.725167\, {\gamma_g}^{-1.485480}\, API^{-0.16471} {T_i}^{-0.091330} {P_b}^{0.047094} \quad \textbf{5-14}$$

Where γ_g and API are the gas specific gravity and oil gravity, respectively

T_i : Pump suction temperature (°F), P_b: Bubble point pressure (psi)

Equation 5-11 is used to determine the reduced solution gas-oil-ratio at any $P_i < P_b$. After calculating the R_{sr}, the solution gas-oil-ratio (R_s), SCF/STB, is determined using Equation 5-15. The R_{sb} (SCF/STB) is the solution gas-oil-ratio at bubble point pressure. To determine the R_s, the

R_{sb} must be calculated. Standing's (1947) correlation was used to determine the solution gas-oil-ratio at bubble point pressure (Equation 5-16).

$$R_s = R_{sr}\, R_{sb} \quad \textbf{5-15}$$

$$R_{sb} = \gamma_g \left[\left(\frac{P_b}{18.2} + 1.4\right) 10^{(0.0125API - 0.00091T_i)}\right]^{1.2048} \quad \textbf{5-16}$$

ii. Determination of Oil Formation Volume Factor (B_o):

The amount of oil at reservoir pressure and temperature required to produce one stock tank barrel of oil is referred to as the oil formation volume factor (B_o). Because oil is slightly compressible, the B_o slightly increases as the reservoir pressure declines due to liquid phase expansion at $P > P_b$. When the $P < P_b$, the $B_o$decreases due to gas phase evolution, as illustrated in Figure 5-8.

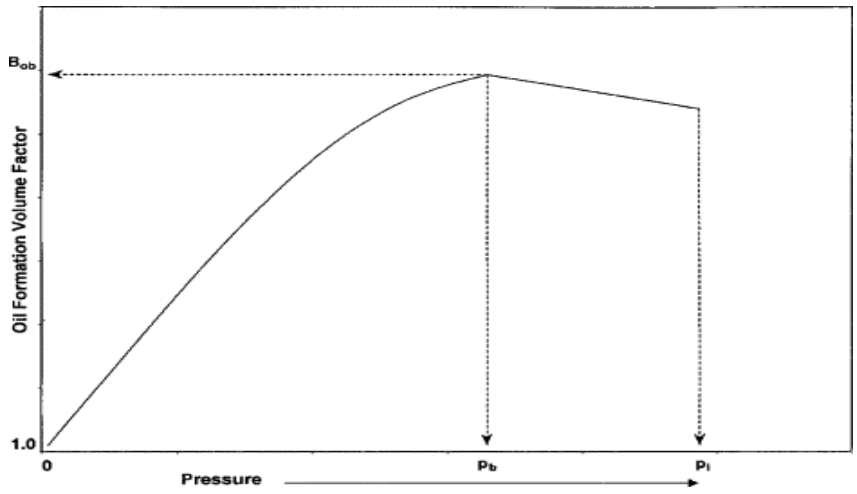

Figure 5-8: Typical Oil Formation Volume Factor Diagram (Ahmed, 2010)

For pressure above the bubble point, $P_i > P_b$, the oil formation volume factor (B_o) is estimated using Standing's (1947) correlation:

$$B_o = B_{ob}\, exp\{-C_o\,(P_i - P_b)\} \quad \textbf{5-17}$$

Where B_{ob} is the formation volume factor at bubble point (RB/STB), and C_o is the oil compressibility (psi^{-1}). The Vazques-Beggs (1980) empirical correlation was used to determine the oil compressibility, C_o :

$$C_o = \frac{-1433 + 5R_S + 17.2T_i - 1810\,\gamma_g + 12.61\,API}{10^5 P_i} \quad \textbf{5-18}$$

Where R_s is equal to R_{si} (SCF/STB) at $P_i > P_b$. For pressures below the bubble point, $P_i < P_b$, the oil formation volume factor (B_o) is estimated using Velarde (1996) correlation:

$$B_o = 1.023761 + 0.000122\left[R_S^{0.413179}\,\gamma_g^{0.210293} API^{0.127123} + 0.019073 T_i\right]^{2.465976} \quad \textbf{5-19}$$

where the solution gas-oil-ratio, R_s, (SCF/STB) is described by Equation 5-15. Thus, the B_o is a function of pressure, temperature, gas solubility, and the bubble point pressure.

iii. Determination of Gas Formation Volume Factor (B_g):

To estimate the free gas formation volume factor (B_g) for the Black oil reservoirs (Equation 5-29), the following parameters must be estimated: the gas compressibility factor (z) using the Brill and Beggs (1991) correlation (Equation 5-20), the pseudocritical properties (P_{pc} and T_{pc}) using the Brown *et al*. (1948) correlation (Equation 5-25 and 5-26), and pseudoreduced properties (P_{pr} and T_{pr}) using Standing and Katz (1942) correlation (Equation 5-27 and 5-28). These equations are as follows:

$$z = A + (1 - A)e^{-B} + C\,P_{pr}^{D} \quad \textbf{5-20}$$

$$A = 1.39\left(T_{pr} - 0.92\right)^{0.5} - 0.36\,T_{pr} - 0.1 \quad \textbf{5-21}$$

$$B = \left(0.62 - 0.23\,T_{pr}\right)P_{pr} + \left[\frac{0.066}{T_{pr} - 0.86} - 0.037\right]{P_{pr}}^2 + \frac{0.32{P_{pr}}^2}{10^{9\,(T_{pr}-1)}} \quad \textbf{5-22}$$

$$C = 0.132 - 0.32 \log_{10}(T_{pr}) \quad \textbf{5-23}$$

$$D = 10^{(0.3106 - 0.49 T_{pr} + 0.1824 T_{pr}^2)} \quad \textbf{5-24}$$

$$P_{pc} = 709.604 - 58.718\, \gamma_g \quad \textbf{5-25}$$

Where P_{pc} in psi. Valid for H_2S < 3%, N_2 < 5%, and total content of inorganic compounds < 7%.

$$T_{pc} = 170.491 + 307.344\, \gamma_g \quad \textbf{5-26}$$

Where T_{pc} in °R. Valid for H_2S < 3%, N_2 < 5%, and total content of inorganic compounds < 7%.

$$P_{pr} = \frac{P_i}{P_{pc}} \quad \textbf{5-27}$$

Where P_i is the pump intake pressure, psi

$$T_{pr} = \frac{T_i}{T_{pc}} \quad \textbf{5-28}$$

Where T_i is the pump suction temperature, °R (°F+460=°R)

$$B_g = 0.0283\, \frac{zT_i}{P_i} \quad \textbf{5-29}$$

Where B_g in ft^3/SCF

5.4 Modelling Approach

This section addresses the machine learning approaches used to fulfil the research goals. The objective is to enable computers to detect patterns within the data and utilize it as a basis for predicting ESP failures. The research aimed to use ESP labelled datasets as inputs to predict multiple specific failure modes (SFM). A supervised multi-class classification model was used. The following sections detail the construction and description of the five modelling designs developed, as presented in Figure 5-9. These designs differ in the input variables that are used to train the models. Next, the ML models constructed for all the designs and validated using the performance measures are detailed. The best model with the highest accuracy in forecasting ESP Failures was then identified. The results and discussions are presented in Chapter 6.

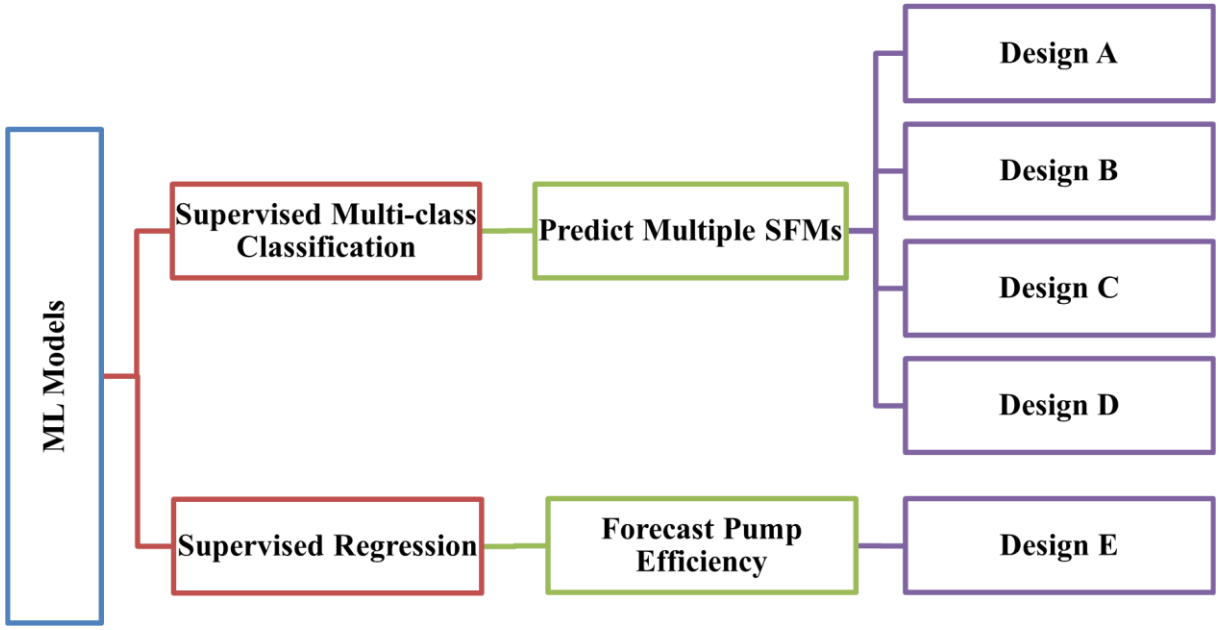

Figure 5-9: The Categories of Machine Learning Designs Constructed

5.4.1 Design A

Design A made use of all the ESP variables listed in Table 3-1. No changes were made in the data before the analysis. All these data give insight into the pump's motor efficiency, wellhead, and surface parameters, as well as the downhole pump's performance.

5.4.2 Design B

In Design B, the objective was to introduce physics into the analytical approach and reduce the number of variables. Several changes were made to the dataset utilized in design A. Pressure and temperature differences were calculated and used to replace the original parameters. Table 5-1 presents the calculated variables for Design B. All the original data were eliminated from the analysis except the motor frequency. The analysis must consider the speed of the motor, which is reflected in motor frequency.

Table 5-1: Design B - Utilized Variables

Calculated Variable	Relationship	Representing
$\Delta \boldsymbol{P}_{\boldsymbol{pump}}$	$P_d - P_i$	Pressure difference in pump, psi
$\Delta \boldsymbol{P}_{\boldsymbol{tubing}}$	$P_d - P_{wh}$	Pressure loss in tubing, psi
$\Delta \boldsymbol{P}_{\boldsymbol{flowline}}$	$P_{wh} - P_{fl}$	Pressure loss in flowline, psi
$\Delta \boldsymbol{P}_{\boldsymbol{ann}}$	$P_i - P_{csg}$	Pressure loss in annulus, psi
OP. Temp	$T_m - T_i$	Operating Temperature, °F
Motor Frequency	From VSD	Motor frequency, Hz

5.4.3 Design C

To get a better understanding of the relationship between η_{pump} and GVF, as well as their influence on ESP failure prediction, Design C was developed. This design is a classification model that predicts the SFM using the GVF and η_{pump} as input variables, among other factors. Design C makes use of all the ESP variables listed in Table 3-1, in addition to η_{pump} (Equation 5-5) and GVF (Equation 5-8). In essence, it adds these two variables to Design A.

5.4.4 Design D

Design D, like Design B, seeks to incorporate physics into the analytical approach and reduce the number of variables compared to Design C. The input variables of this design are indicated in Table 5-2. They include the variables from Design B, coupled with the computed pump efficiency (Equation 5-5), and GVF (Equation 5-8) to serve as this model's input variables.

Table 5-2: Design D-Input Variables

Input Variables
ΔP_{pump}
ΔP_{tubing}
$\Delta P_{flowline}$
ΔP_{ann}
OP. Temp
Input Electrical Power
Motor Frequency
GVF %
Pump Efficiency %

5.4.5 Design E

In Design E, the ESP-labelled datasets were utilized as inputs, and a supervised regression model was employed to forecast the pump efficiency (η_{pump}). Daily production and ESP data were used to estimate the fluid properties, as discussed in Section 5.3.1, and determine the pump efficiency (η_{pump}). The estimated GVF was combined with the production and ESP data to provide the input for the regression model, as shown in Table 5-3. Next, several ML models were constructed and validated using the regression performance measures. The best model with the highest accuracy to predict the output variable (pump efficiency) was identified.

Table 5-3: Design E-Input Variables

Variables	Description
MOT_FREQ	Motor frequency, Hz
$\Delta \boldsymbol{P_{pump}}$	Pressure difference in pump, psi
$\Delta \boldsymbol{P_{tubing}}$	Pressure loss in tubing, psi
$\Delta \boldsymbol{P_{flowline}}$	Pressure loss in flowline, psi
$\Delta \boldsymbol{P_{ann}}$	Pressure loss in annulus, psi
OP. Temp	Operating Temperature, °F
I_MOTOR	Motor current, Amps
V_MOTOR	Motor voltage, Volts
GVF	Gas Volumetric fraction, %
$\boldsymbol{Q_o}$ **[bbl/d]**	Oil production rate, BPD, low frequency data (daily)

5.5 Prediction Periods Determination

The goal of this study was to predict the specific failure modes of the ESP before they occur. The research tested several prediction periods to train the classifier and predict the given SFM's before they happen. The prediction period is defined as the time duration before the failure occurs when the model predicts the failure. It is reasonable to assume that forecasting true failure becomes more difficult as the prediction period lengthens (Khabibullin *et al*., 2020). In order to avoid over or under engineering, the selection of the right prediction period is crucial. Early repair

and maintenance may help minimize downtime and production losses due to ESP failure by early diagnosis and implementing suitable action plans. Longer prediction periods, however, may lead to excessive maintenance costs and resources (Takacs, 2017), and lower reliability of the failure predictions due to higher rates of false alarms. To use the machine learning models in the most efficient way, one must first find the optimal prediction period.

The ideal prediction period depends on the type of the ESP failure. In some cases, little to no early signs may be observed in the data prior to the failure. In some other cases, the symptoms can be captured in the data long before the failure is detected at the surface, as summarised in Table 3-7. These early warning indicators may show that the system is losing its efficiency. Therefore, ML models are used to aid in distinguishing the actual ESP failures from false alarms. A field engineer's judgment on the type of failure is the final arbiter of the reasonable time frame for early failure prediction, based on the field history and the availability of resources.

In this study, all designs were used with prediction periods (PP) of three hours, one day, and three days. In addition, durations of 5 and 7 days were examined only for design A and analyzed, as detailed in Section 6.1.1.

5.6 ML Algorithms

Several ML models were constructed for each design and validated using the performance measures. The best model with the highest accuracy to forecast the ESP Failure was identified. The supervised classification and regression models investigated are as follows (Pedregosa *et al*., 2011):

1. Random Forest (RF): The RF employs several decision trees on distinct subsets of a given dataset and averages them to enhance the dataset's prediction accuracy.

2. Decision Tree (DT): In DT, the dataset is divided into groups depending on the values of each attribute. The “Decision Node” and the “Leaf Node” are the two nodes. Decision nodes are used to make decisions and have many branches, while Leaf nodes result from such decisions and have no more branches.
3. Multilayer Perceptron (MLP) Neural Network: The MLP is a deep learning artificial neural network that generates outputs from a set of inputs. Multiple layers of input nodes are linked as a directed graph between the input and output layers. MLP uses backward propagation of errors to train the network.
4. K-Nearest Neighbor (KNN): Despite its simplicity, the K-neighbors is the most extensively used and straightforward technique. The KNN algorithm is a data categorization approach that calculates the likelihood that a data point will belong to one of two groups based on the involvement of the data points closest to it. The appropriate K value is a user-defined constant, strongly depending on the data.
5. Multinomial Logistic Regression (LR): The LR is a classification technique that assigns data to a set number of classes. Probability values are generated from its output using the logistic sigmoid function.
6. Gaussian Naive Bayesian Classification (GNB): In the case of GNB, each class is assumed to have a normal distribution. Predictive modeling may benefit significantly from this approach, despite its simplicity.
7. Support Vector Regression (SVR): Similar to the support vector machine (SVM), the SVR predicts discrete values. It is the primary goal of SVR to discover the hyperplane with the greatest number of points.

8. Least Absolute Shrinkage and Selection Operator (LASSO): LASSO is a linear regression approach that employs variable selection and regularization to increase the accuracy and interpretability of the prediction. It employs L1 regularization, which penalizes coefficients in accordance with their absolute magnitude.
9. Ridge Regression: Ridge method is a way to create a parsimonious model when the number of predictor variables in a set exceeds the number of observations, or when a data set has multicollinearity. It employs L2 regularization, which applies an L2 penalty equal to the square of coefficient magnitude.
10. Polynomial Regression: In this method, the relationship between the independent variable x and the dependent variable y is modeled as an n^{th}-degree polynomial in x.

5.7 ML Performance Metrics

Performance metrics are critical components of every machine learning model. They aid in evaluating the model's performance and monitoring the system (Grandini and Visani, 2020). This section discusses the performance metrics used to evaluate multi-class classification and regression models. The results of this section are presented and discussed in Chapter 6.

5.7.1 Multi-class Classification Metrics

The objective is to develop and choose a highly accurate model to predict ESP specific failure modes (SFM's). Several classification metrics were used to analyze the model and evaluate its performance. The classifications metrics are as follows:

1. Normalized Confusion Matrix: The normalized confusion matrix was constructed to analyze and illustrate the outcomes of the machine learning model. As illustrated in

Figure 5-10, each column represents the model's SFM predictions, while each row represents the actual SFM classifications in the data. Dividing a row element by the total number of the whole row produces the normalized confusion matrix. Therefore, the final normalized matrix indicates the class making the best forecast for a true label.

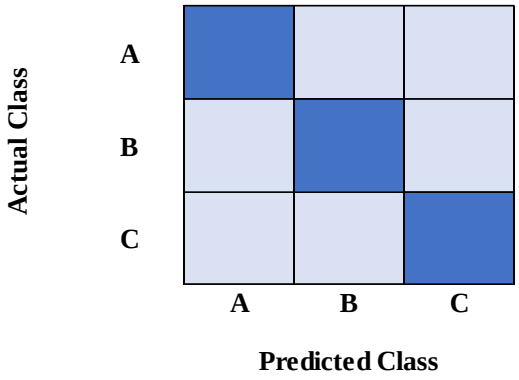

Figure 5-10: Multi-class Classification Confusion Matrix

2. Precision: Precision indicates the proportion of correct predictions genuinely belong to a specific class. A model with a high precision does not produce too many false alarms. Table 5-4 shows the numerical definitions for all the parameters.
3. Recall: Recall expresses the ratio of correct positive predictions produced from all positive instances in a dataset. A model with a high recall, does not miss many positive occurrences
4. F1 Score: The F1 Score is the mean of precision and recall. It accounts for both false positives and false negatives. As a result, it performs well on an imbalanced dataset. When both Precision and Recall are low, a model has a low F1 score, and a high F1 score is achieved when both Precision and Recall are high.

Table 5-4: Typical Classification Metrics Description (Pedregosa *et al.*, 2011)

Metrics	Descriptions
Precision	= TP / (TP + FP), a trade-off value to Recall.
Recall	= TP / (TP + FN), a trade-off value to Precision.
F1 Score	= 2TP / (2TP + FP + FN), a balanced metric between Precision and Recall.
*TP: True Positive, TN: True Negative, FP: False Positive, and FN: False Negative	

5. Cohen's Kappa Statistic (κ): Cohen's kappa is a statistical measure of inter-rater agreement for qualitative items (McHugh, 2012). The performance of the classifier is determined by evaluating the level of agreement between the predicted and actual values. It is often regarded as a more robust measure than simple accuracy computation due to its capacity to manage multi-class and imbalanced class classifications. The interpretation of the Kappa result is given in Table 6-7.

Table 5-5: Kappa's Result Interpretations

Kappa Results	Indication
values ≤ 0	No agreement
0–0.20	None to minor agreement
0.21–0.40	Moderate agreement
0.41–0.60	Significant agreement
0.61–0.80	Substantial agreement
0.81–1.00	Almost perfect agreement

5.7.2 Regression Metrics

The objective of Design E is to develop and choose a highly accurate model to forecast the η_{pump} (Section 5.4.5). Several regression metrics were used to analyze the model and evaluate its performance. The regression metrics are summarized in Table 5-6. All these metrics will be shown in the analysis of Design E in Chapter 6.

Table 5-6: Typical Regression Metrics Description (Pedregosa *et al.*, 2011)

Metrics	Description
RMSE	Root Mean Squared Error of the test dataset (a lower value is better).
MAE	Mean Absolute Error of the test dataset (a lower value is better).
R^2	R^2 of the test dataset (a higher value is better).
EV Test	Explained Variance of the test dataset (a higher value is better).

Chapter 6 Results and Discussion

The results of the ESP multi-failure ML models are investigated in this chapter, using various machine learning models across the three-hour to 7-day prediction periods. This chapter summarizes and discusses the findings of the five modelling designs described in Chapter 5.

6.1 Design A

Design A is intended to use all the raw ESP variables that provide insight into the pump's motor efficiency, surface parameters, and downhole pump parameters. As described in Data Pre-processing (Section 5.2), GridSearchCV was used to optimize hyperparameters to improve the performance of each ML model. GridSearchCV employs both Grid Search and K-Fold cross-validation (5 folds) to tune the model's hyperparameters. These parameters are then used to construct a new ML model. Table 6-1 details the tuned parameters for each ML model in Design A.

Table 6-1: Design A-Hyperparameters Optimization Results

ML Models	Hyperparameter to be Tuned			Tuned Hyperparameter
	Hyperparameters		Different Values	
RF	max_samples	Maximum number of samples in each tree	[1000, 2000, 3000, 4000, 5000]	5000
	n_estimators	Number of trees in the forest	[200, 500, 1000, 2000]	1000
	max_depth	The longest path between the root node and the leaf node	[3,4,5,6]	6
	min_samples_split	Min number of data points placed in a node before the node is split	[5,10,15]	5
DT	max_leaf_nodes	Maximum number of leaf nodes a decision tree can have	[1, 2, 3, 4, 5, 6, 7, 8, 9, 10]	10
	min_samples_split	Minimum number of samples required to split an internal node	[9, 10]	9
MLP	hidden_layer_sizes	Number of layers and nodes in the Neural Network (NN)	[(50,50), (50,100), (100,1)]	(50,100)
	Activation Function	It defines how the weighted sum of the input is transformed into an output from a node(s) in a layer of NN	[relu, tanh, logistic]	relu
	Alpha	It combats overfitting by constraining the size of the weights	[0.0001, 0.05]	0.0001
	learning_rate	It controls the weights of NN	[constant, adaptive]	adaptive
	Solver	The solver iterates until convergence	Adam	Adam
KNN	K values	The count of the nearest neighbours	From 1 to 100	10
GNB	var_smoothing	Portion of the biggest variation among all variables that is added to other variances to ensure computation stability.	From 1 to 10^{-9}	1.0
LR	C	Penalty to reduce overfitting	[0.001,0.01, 0.1, 1.0, 10]	0.001

Figure 6-2 summarizes a comparison of average F1-score results of the tested ML models for Design A. In this design, K-Nearest Neighbor (K=10) is the best model for predicting ESP failures over all the prediction periods in terms of precision, recall, and consequently the F1-score.The accuracy, precision and recall results for each SFM are compared for various K values in Figure B- 4, Figure B- 5, Figure B- 6, Figure B- 7, and Figure B- 8 (appendix B). It is reasonable to conclude that K=10 provides the maximum precision, recall, and F1-score.

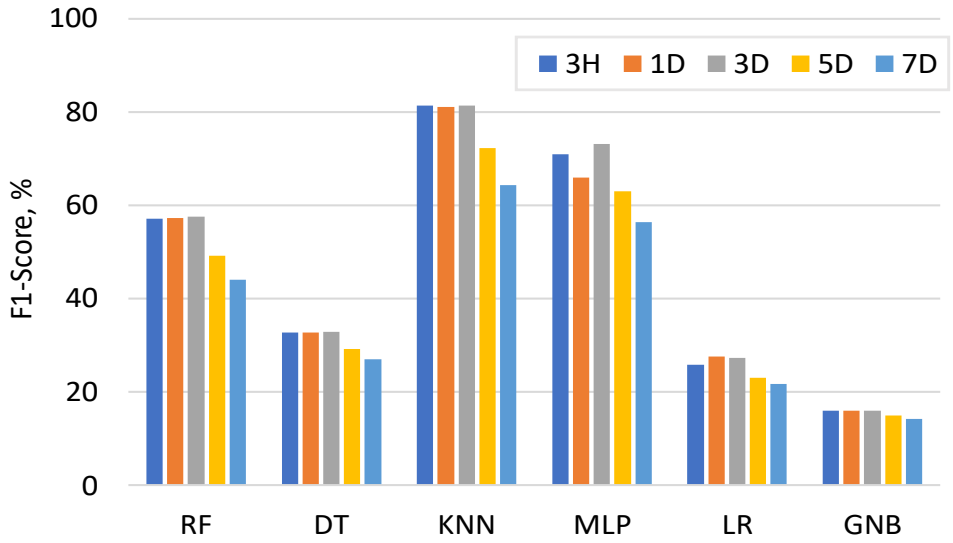

Figure 6-1: Design A- Summary of Model Performance (F1-Score)

When dealing with multi-class classification for imbalanced classes, Cohen's Kappa Coefficient is more instructive than the overall accuracy. Figure 6-2 summarizes a comparison of Cohen's Kappa Coefficients (κ) of the tested ML models for Design A. The KNN model achieves the highest Kappa (κ) among the other models for all the PP's (κ ~ 0.7-0.86), indicating substantial to almost-perfect agreements between the actual and predicted SFMs, as described in Table 5-5.

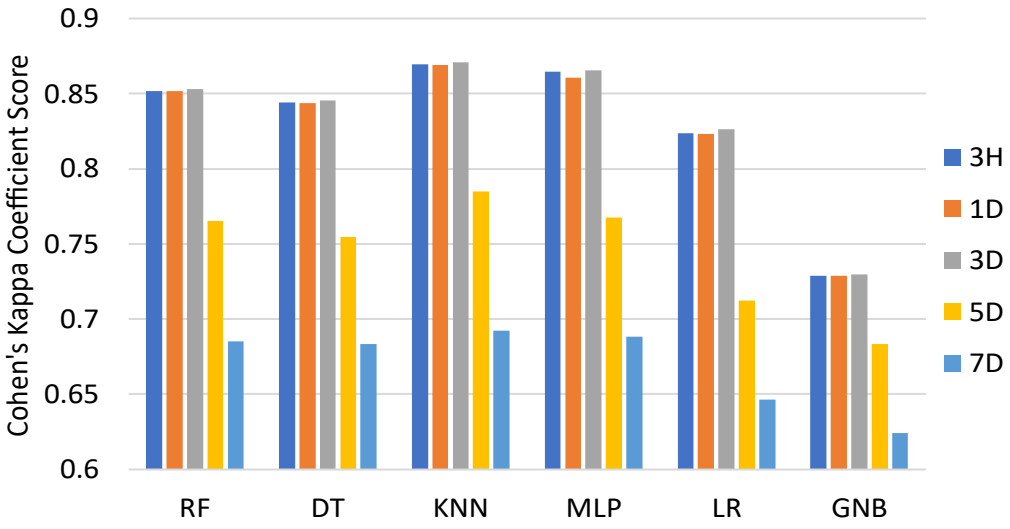

Figure 6-2: Design A- Summary of Cohen's Kappa Coefficient (κ) Results

Table 6-2 highlights the performance of Design A's best model, KNN, for the tested prediction periods and for each ESP specific failure. As observed, the KNN model demonstrates a high degree of precision and recall across multiple SFMs, indicating that the algorithm is performing effectively.

Table 6-2: Design A-KNN Results over All Prediction Periods

	Prediction Period of 3 hours			Prediction Period of 1 Day			Prediction Period of 3 Days			Prediction Period of 5 Days			Prediction Period of 7 Days		
SFM	**Precision**	**Recall**	**F1-score**	**precision**	**Recall**	**F1-score**	**precision**	**Recall**	**F1-score**	**precision**	**Recall**	**F1-score**	**precision**	**Recall**	**F1-score**
Broken Shaft	1	0.21	0.35	1	0.21	0.35	1	0.21	0.35	0.89	0.17	0.28	0.83	0.10	0.18
MTR Cable	0.95	0.97	0.96	0.94	0.97	0.96	0.96	0.97	0.97	0.77	0.84	0.8	0.68	0.77	0.72
High Tm	0.61	0.74	0.67	0.62	0.68	0.65	0.59	0.71	0.64	0.58	0.67	0.62	0.58	0.54	0.56
Plugged with sand	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.87	0.89	0.88	0.85	0.72	0.78
Unknown SFM	0.77	1	0.87	0.77	1	0.87	0.77	1	0.87	0.67	0.87	0.76	0.68	0.69	0.68
Overload	0.99	0.75	0.85	0.99	0.75	0.85	0.99	0.75	0.85	0.86	0.69	0.77	0.84	0.61	0.71
Power Fail	0.98	0.62	0.76	0.98	0.62	0.76	0.98	0.62	0.76	0.91	0.54	0.69	0.78	0.49	0.60
Sensor Failure	0.94	0.64	0.76	0.94	0.64	0.76	0.94	0.64	0.76	0.83	0.55	0.66	0.69	0.46	0.55
Under Voltage	0.93	0.96	0.95	0.93	0.96	0.95	0.93	0.96	0.95	0.86	0.87	0.86	0.81	0.70	0.75
Underload	0.77	0.84	0.8	0.77	0.81	0.79	0.8	0.83	0.82	0.67	0.68	0.67	0.68	0.53	0.59
Normal	1	0.98	0.99	1	0.98	0.99	1	0.98	0.99	0.99	0.95	0.96	0.99	0.93	0.96
Cohen's Kappa Coefficient Score	0.8696			0.8691			0.8708			0.7850			0.6925		

6.1.1 Prediction Period Effects

As indicated before, the KNN algorithm is the best ML model in terms of precision, recall, and the F1-score for predicting ESP Failure over all prediction periods. Figure 6-3 depicts the F1-score for the KNN model (Design A) for each ESP specific failure mode across the five tested prediction periods. The F1-score decreases as the duration of the forecast period increases, particularly for the intervals of 5 and 7 days. In addition, as seen in Figure 6-2, the Cohen's Kappa Coefficient Score for the KNN model (Design A) decreases when the prediction period extends past 3 days.

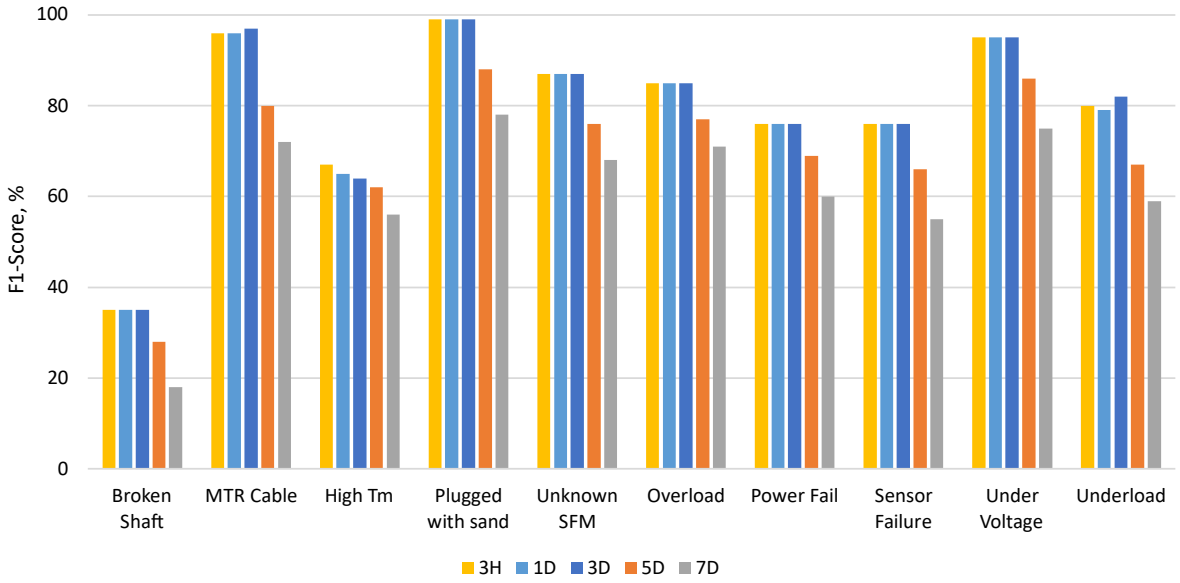

Figure 6-3: Design A- F1-Score for Every SFM Across All PP

Figure 6-4 shows the KNN model's normalized confusion matrix for a three-day prediction period. The matrices for the prediction periods of three hours and one day are also shown in appendix B, Figure B- 1 and Figure B- 2. The True positive rate is the highest for the cases shown in dark blue. The matrices demonstrate that the model correctly predicts most ESP specific failures, while exhibiting a low rate of false alarms. The SFM with the worst prediction performance is "Broken Shaft", which is correctly predicted only 21% of the overall count. In all other cases, the

broken shaft failure is predicted by the model, but the cause is predicted as “Unknown SFM”. As discussed in Section 3.3.3, the lack of vibration parameters makes it difficult, based on the available data in this study, to recognize the symptoms preceding broken shaft failures.

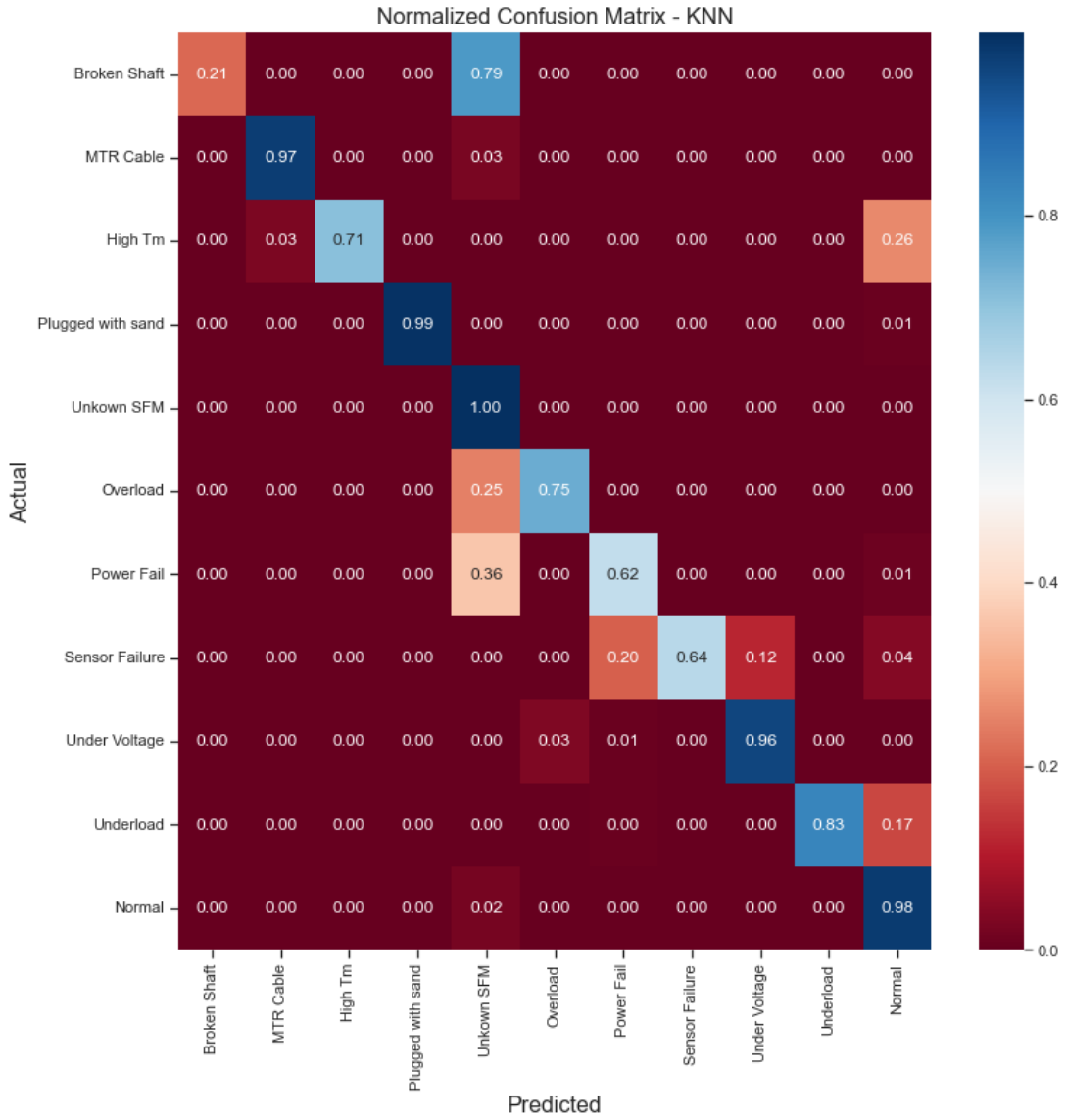

Figure 6-4: Design A-KNN Normalized Confusion Matrix (3 Days PP)

Figure 6-5 depicts the KNN normalized confusion matrix with a 7-day prediction period. As shown in Figure 6-5, the model successfully predicts only three ESP specific failures (70-77%) with moderate average F1-score of 64 % and κ of 0.69, while exhibiting a greater rate of false alarms than the previously presented prediction periods, as shown in Table 6-2. Figure B- 3

displays slightly better, yet similar results for the KNN normalized confusion matrix for a 5-day prediction period with average F1-score of 70% and κ of 0.78. These findings align with the statement of Khabibullin *et al*. (2020) that forecasting true failure becomes more difficult as the prediction period lengthens. For the next designs, only the better performing prediction periods (three hours, one day, and three days) are included in the analysis to avoid lower reliability of failure predictions due to increasing false alarm rates.

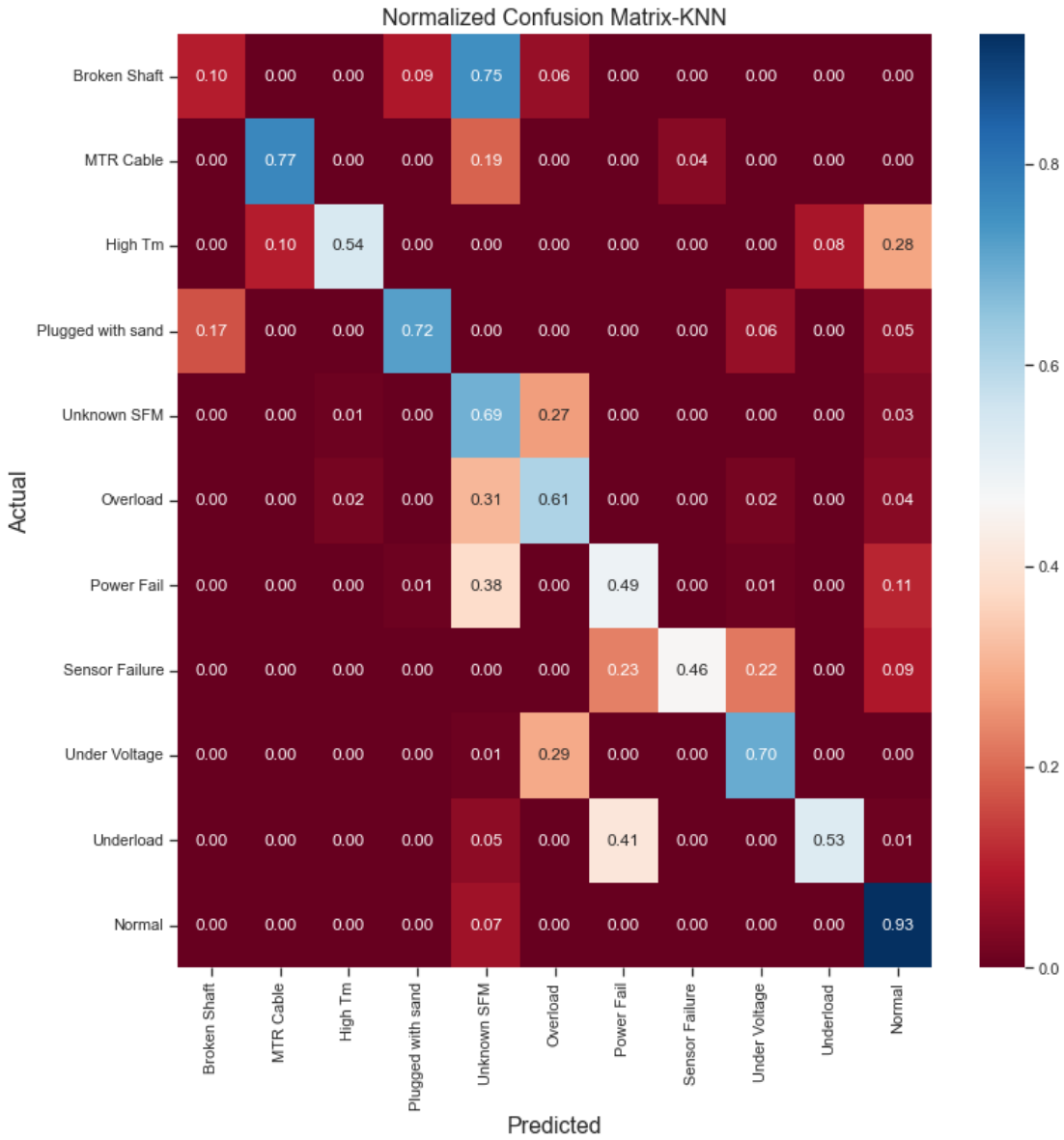

Figure 6-5: Design A-KNN Normalized Confusion Matrix (7 Days PP)

6.1.2 Variables Significance Evaluation

In machine learning, feature or variable importance gives a score to input variables based on how well they predict a target variable. It provides information on the data, the model, and the feature selection that may increase the efficiency and effectiveness of ML prediction models. The Random Forest classifier is a typical method for quantifying variable importance; consequently, it was used to estimate and illustrate the variable importance.

According to the variable importance presented in Figure 6-6, five variables have the most significant influence on predicting failures: motor current, motor frequency, wellhead pressure, flowline pressure, and motor voltage. Motor current and frequency are the most effective indicators for forecasting ESP specific failures with a total of 62 % influence on forecasting failures. This finding is consistent with the observation that electrical failure is the most common cause of operational failures in this study, making motor current and voltage the most impacted variables.

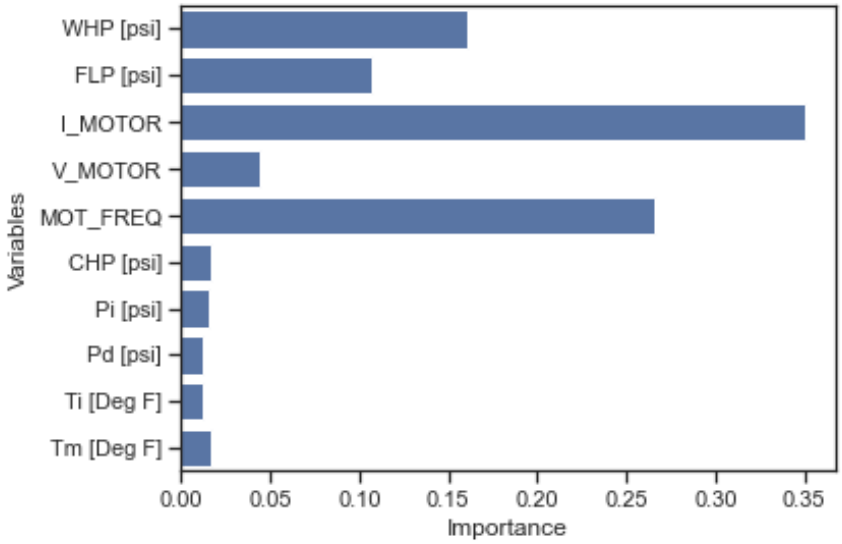

Figure 6-6: Design A-Variables Importance

The five mentioned variables were examined in greater detail to ascertain their relationship, as shown by a correlation matrix in Figure 6-7. Figure 6-7 depicts a significant positive correlation

for wellhead pressure with the flowline pressure and motor frequency. A variable speed drive (VSD) is often used to control the rotation speed of a pump motor and mitigate pressure fluctuations at the wellhead. Figure 6-7 also demonstrates a strong positive correlation between the motor voltage and frequency. Under a constant load (torque), the motor's speed is proportional to the supply voltage (Takacs, 2017).

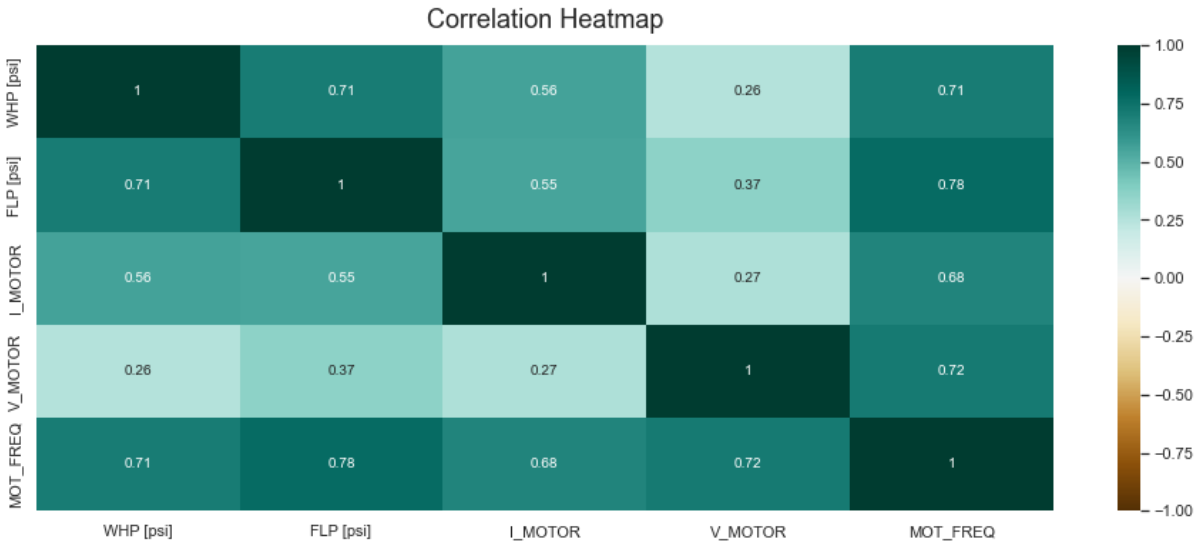

Figure 6-7: ESP Variables Correlation Heatmap Matrix-Selected Variables

6.2 Design B

In Design B, pressure and temperature differences are calculated and used as the correlating variables, together with motor frequency. The objective is to reduce the number of data and keep only the more significant variables. Table C- 1 details the tuned parameters for each ML model in Design B. K-Nearest Neighbor (K=10) is the best model with the greatest F1-score for predicting ESP Failures across all SFMs and throughout all prediction periods. Figure 6-8 summarizes a comparison of average F1-scores for the tested ML models for Design B. The K value is chosen based on the outcomes of Figure B- 11, Figure B- 12, and Figure B- 13, with K=10 providing the highest precision, recall, and F1-score.

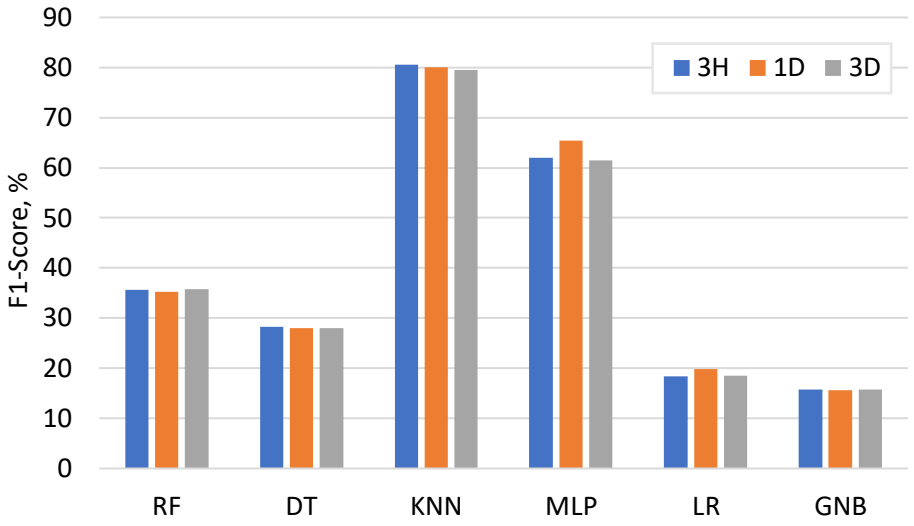

Figure 6-8: Design B- Summary of Model Performance (F1-Score)

Figure 6-9 summarizes a comparison of Cohen's Kappa Coefficient (κ) results of the tested ML models for Design B. The KNN model achieves the highest Kappa (κ) among all the models, corresponding to an almost perfect agreement between the actual and predicted SFMs (> 0.86).

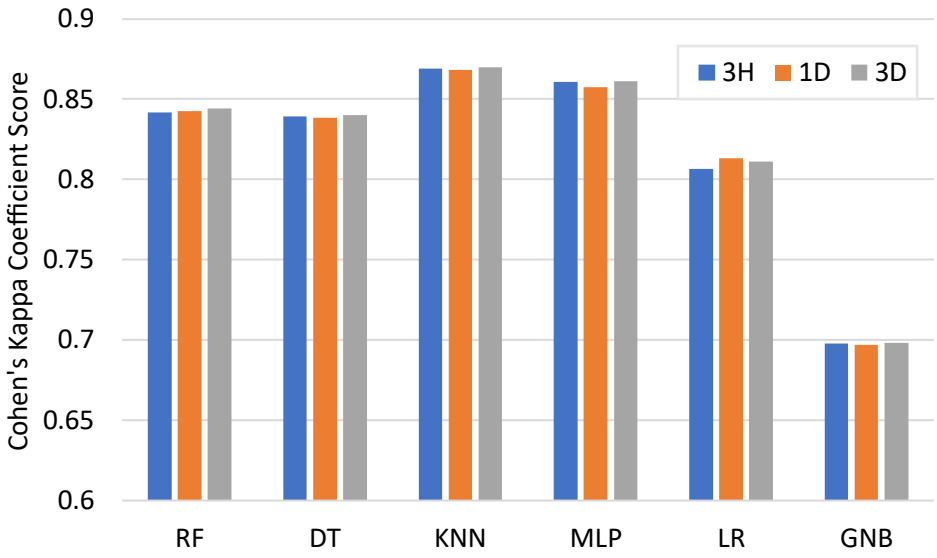

Figure 6-9: Design B- Summary of Cohen's Kappa Coefficient (κ) Results

Due to their unsatisfactory performance in this study, Gaussian Naive Bayes (GNB) and Multinomial Logistic Regression (LR) will not be employed in Designs C and D. While GNB is

straightforward and simple to construct, its lower performance may be the result of the unrealistic assumption that all features are independent, limiting its practical use. Moreover, it is challenging to simplify complicated relationships when using LR since this approach thrives when the dataset is linearly separable.

Table 6-3 describes the performance of Design B's best model, KNN, for all the prediction periods and for each ESP specific failure. The KNN model demonstrates a high degree of precision and recall across multiple SFMs, indicating that the algorithm is performing effectively. Overall, design B can provide a failure prediction accuracy, almost as good as design A, without the need to use many variables.

According to the variable importance plot in Figure 6-10, motor frequency has the most significant impact on predicting failures (66%) in Design B. This finding is consistent with the variable importance of Design A. The $\Delta P_{flowline}$ is the second most significant variable, as it is physically a strong function of flowrate produced through the pump. The correlation heatmap in Figure B- 14 clearly illustrates that ΔP_{tubing} and ΔP_{pump} have a high positive correlation with motor frequency (65-80%).

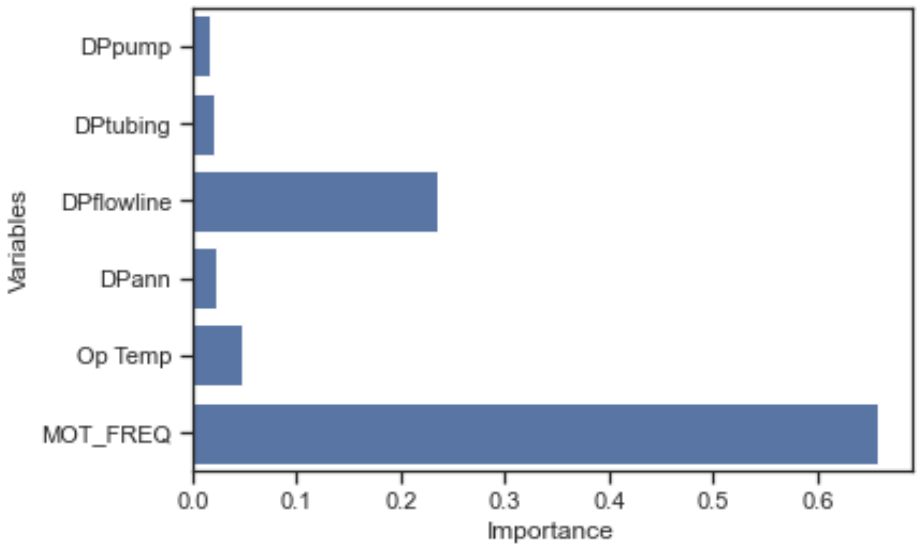

Figure 6-10: Design B-Variables Importance

Table 6-3: Design B-KNN Results over Three Prediction Periods

	Prediction Period of 3 hours			Prediction Period of 1 Day			Prediction Period of 3 Days		
SFM	**Precision**	**Recall**	**F1-score**	**precision**	**recall**	**F1-score**	**precision**	**recall**	**F1-score**
Broken Shaft	1	0.21	0.35	1	0.21	0.35	1	0.21	0.35
MTR Cable	0.97	0.94	0.96	0.96	0.94	0.95	0.97	0.94	0.96
High Tm	0.58	0.67	0.62	0.6	0.59	0.6	0.61	0.54	0.57
Plugged with sand	0.99	0.97	0.98	0.99	0.97	0.98	0.99	0.97	0.98
Unknown SFM	0.77	1	0.87	0.77	1	0.87	0.77	1	0.87
Overload	0.97	0.74	0.84	0.96	0.74	0.84	0.97	0.73	0.84
Power Fail	0.96	0.61	0.75	0.95	0.61	0.74	0.96	0.61	0.74
Sensor Failure	1	0.6	0.75	1	0.6	0.75	0.93	0.56	0.7
Under Voltage	0.92	0.95	0.94	0.92	0.96	0.94	0.92	0.96	0.94
Underload	0.78	0.85	0.82	0.78	0.83	0.8	0.83	0.78	0.81
Normal	1	0.98	0.99	1	0.98	0.99	1	0.98	0.99
Cohen's Kappa Coefficient Score=	0.8689			0.8681			0.8697		

The KNN normalized confusion matrix shown in Figure 6-11 shows the performance of the Design B for a three-day prediction period. Figure B- 10 and Figure B- 9 show similar plots for three-hour and one-day prediction periods, respectively. The results indicate that the model predicts each ESP failure accurately, while exhibiting a low percentage of false alarms. Each value shows the rate of true positives the model has accurately identified based on its recall measure.

Similar to design A, the broken shaft failure is poorly identified owing to the absence of vibration data. Moreover, this design incorrectly classified two SFMs: High Tm and Sensor Failure. Based on Table 3-7, we can infer that the motor current and motor voltage are the primary

early warning signals indicators for these two SFMs, hence their possible absence from the Design B input variables might be the cause of this issue.

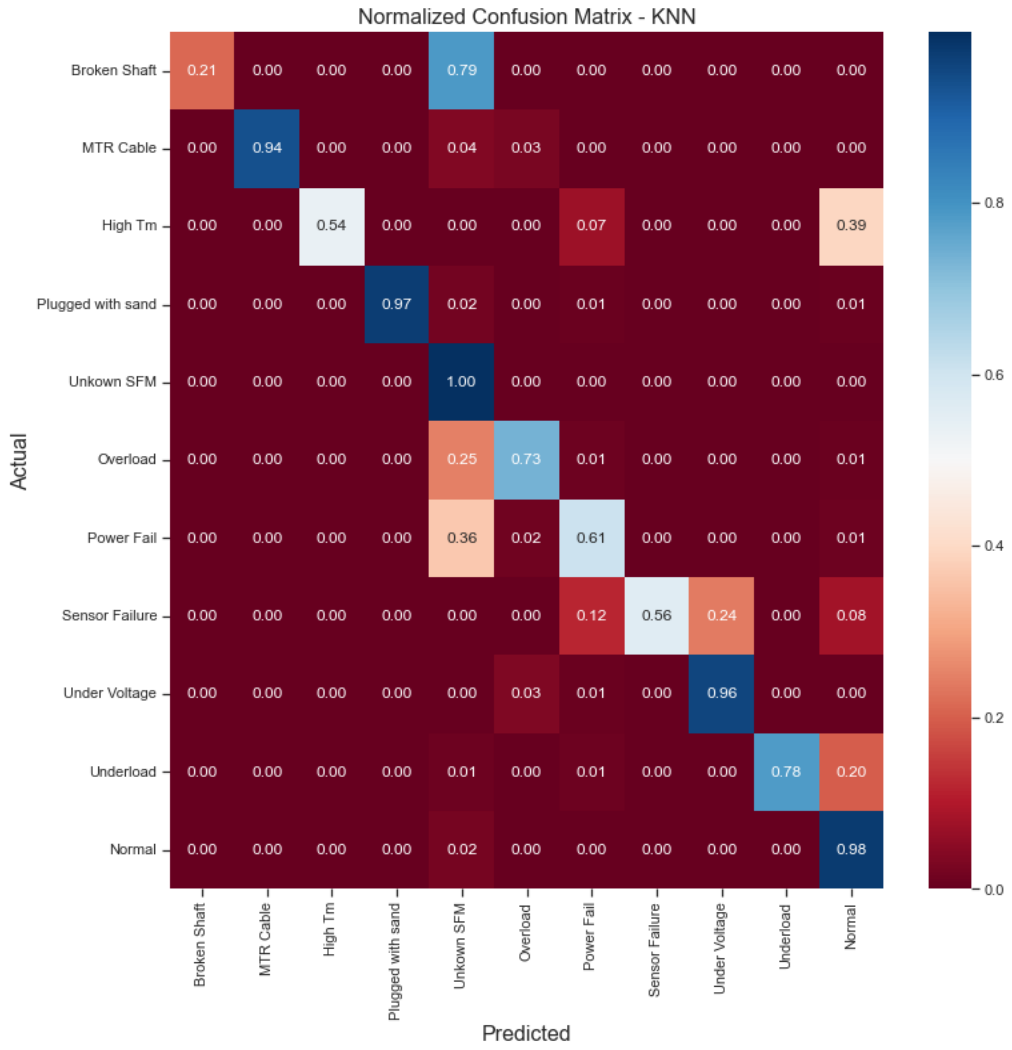

Figure 6-11: Design B-KNN Normalized Confusion Matrix (3 Days PP)

6.3 Design C

Design C is a classification model that predicts the SFM's across the three prediction periods using the GVF and η_{pump} as input variables, added to the other factors in Design A. This design makes use of all the ESP variables listed in Table 3-1, and includes η_{pump} and GVF.

Using boxplots, Figure 6-12 offers a statistical overview of the ranges for the pump efficiency, GVF, and Turpin correlation parameter (ϕ) data. As observed, the GVF values are less

than 50%, indicating that the influence of free gas at the intake conditions for the data in this study is small. In addition, ϕ values are smaller than unity, supporting the observation of the small influence of the free gas at intake conditions, as stable ESP operation is anticipated when (ϕ <1), as described in Section 2.4.2.

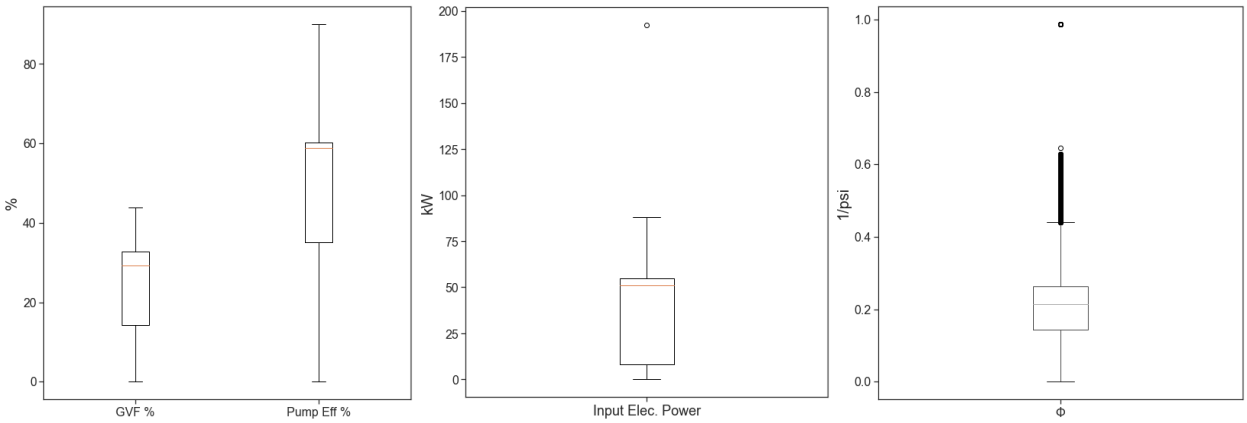

Figure 6-12: Boxplots of the Computed Variables

Figure 6-13 is a scatter plot matrix illustrating the relationships between GVF and pump efficiency with a few of the flow parameters. It reveals a positive linear relationship with the Turpin parameter (ϕ) and a negative linear relationship with pump intake pressure (P_i) for GVF. Both the wellhead pressure and the pump discharge pressure are directly related to the pump's efficiency.

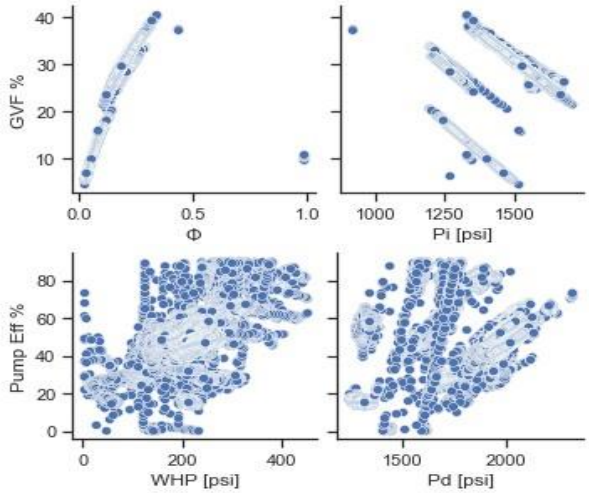

Figure 6-13: Scatter Plots of the Computed Variables

For a qualitative analysis of the relationship between the ESP variables, a correlation matrix was developed, as shown in Figure 6-14. Figure 6-14 illustrates a statistically significant positive correlation between pump efficiency and pump discharge pressure, motor frequency, and wellhead pressure. This result agrees with the concept that the pump efficiency is a function of the production rate, discharge pressure, intake pressure, motor current, and motor voltage (Equation 5-8). Additionally, reducing wellhead pressure and raising discharge pressure are results of the increased fluid velocity in the tubing and a higher production rate and enhanced efficiency.

A considerable inverse correlation is observed between the pump efficiency and motor voltage, GVF, and ϕ. The higher the GVF, the greater the amount of free gas entering the pump intake ($P_i < P_b$), which in turn reduces the pump efficiency owing to gas interference. The GVF has a significant positive association with the motor and pump suction temperatures. Higher GVF results in the greater amount of gas entering the pump and increases the motor temperature owing to the lack of liquid cooling the motor. In addition, the Turpin correlation parameter (ϕ) is derived from the GVF computation (Equation 5-9), indicating that ϕ is proportional to GVF. Moreover, a significant negative correlation is observed between GVF and pump intake pressure, because when the pump intake pressure falls below the bubble point pressure, more gas will develop at the intake. There are no variables with zero correlation, hence all variables are correlated to some degree.

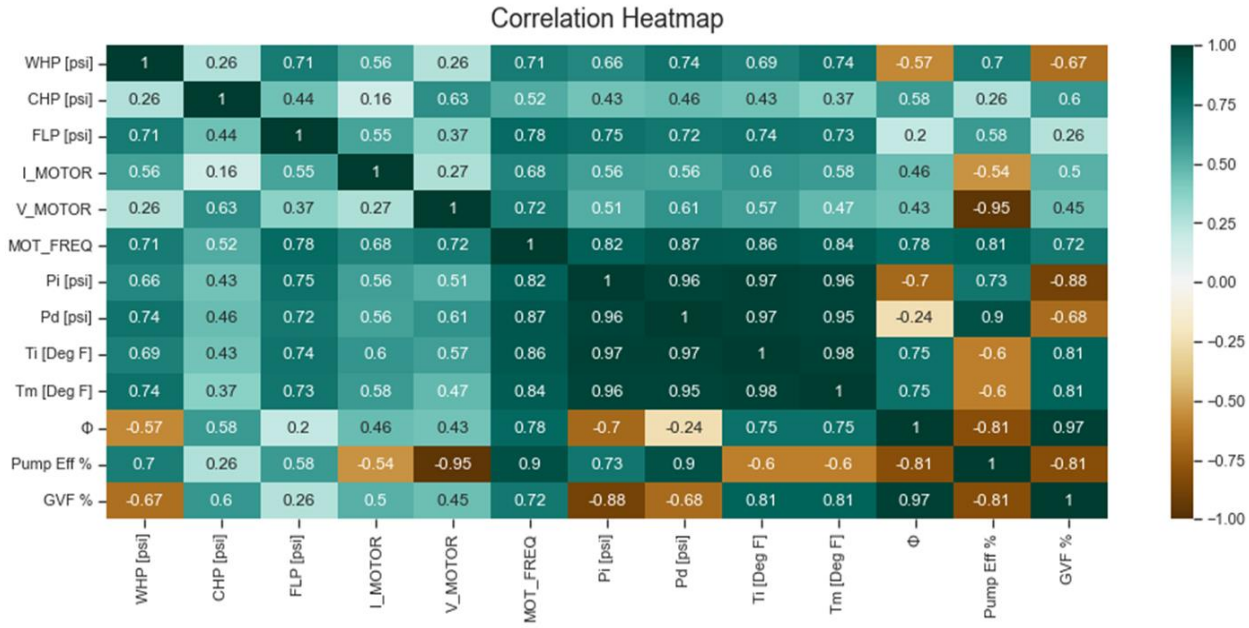

Figure 6-14: Computed Variables Correlation Heatmap Matrix

K-Nearest Neighbor (K=10) is the best model in terms of precision, recall, and the F1-score for predicting the ESP failures over all prediction periods in Design C. Figure 6-15 provides a comparison of Design C's average F1-scores for the tested ML models. Table C- 3 details the tuned parameters for each ML model in Design C. The K value is based on the outcomes of Figure B- 15, Figure B- 16, and Figure B- 17 with K=10 providing the highest precision, recall, and F1-score.

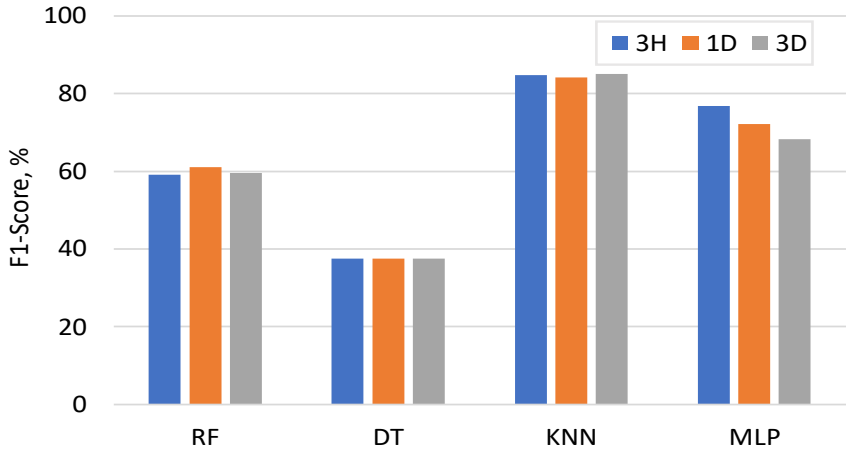

Figure 6-15: Design C- Summary of Model Performance (F1-Score)

Figure 6-16 summarizes a comparison of Cohen's Kappa Coefficient (κ) results for the tested ML models. The KNN model achieves the highest Kappa (κ) among the other models, which corresponds to an almost perfect agreement between the actual and predicted SFM's (> 0.87).

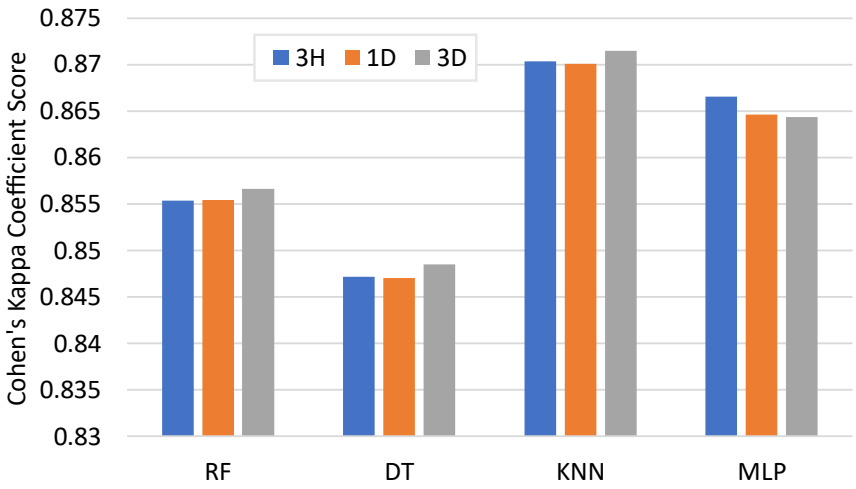

Figure 6-16: Design C- Summary of Cohen's Kappa Coefficient (κ) Results

Table 6-4 highlights the performance of Design C's best model, KNN, across the three prediction periods and for each ESP specific failure. As observed, high degrees of precision and recall are achieved for most SFM's, indicating that the algorithm is performing effectively.

Table 6-4: Design C-KNN Results over Three Prediction Periods

	Prediction Period of 3 hours			Prediction Period of 1 Day			Prediction Period of 3 Days		
SFM	**Precision**	**Recall**	**F1-score**	**precision**	**Recall**	**F1-score**	**precision**	**Recall**	**F1-score**
Broken Shaft	1	0.21	0.35	1	0.21	0.35	1	0.21	0.35
MTR Cable	0.97	0.96	0.97	0.97	0.94	0.96	0.99	0.99	0.99
High Tm	1	0.94	0.97	0.94	0.93	0.93	1	0.96	0.98
Plugged with sand	0.99	1	0.99	0.98	0.99	0.99	0.98	1	0.99
Unknown SFM	0.77	1	0.87	0.77	1	0.87	0.77	1	0.87
Overload	0.98	0.74	0.84	0.98	0.74	0.84	0.99	0.74	0.85
Power Fail	0.98	0.6	0.74	0.98	0.61	0.75	0.98	0.61	0.75
Sensor Failure	0.79	0.6	0.68	0.78	0.56	0.65	0.78	0.56	0.65
Under Voltage	0.92	0.92	0.92	0.93	0.93	0.93	0.92	0.93	0.93
Underload	1	1	1	1	1	1	1	1	1
Normal	1	0.98	0.99	1	0.98	0.99	1	0.98	0.99
Cohen's Kappa Coefficient Score=	0.8704			0.8701			0.8715		

The KNN normalized confusion matrix in Figure 6-17 shows the performance of Design C with a three-day prediction period. Figure B- 19 and Figure B- 18 show similar figures for three-hour and one-day prediction periods, respectively. These plots demonstrate that the model correctly predicts most ESP specific failures, while exhibiting a low rate of false alarms. Design C is superior to Designs A and B in terms of precision, recall, and F1-score, particularly in predicting the high motor temperature and motor underload ESP specific failure modes.

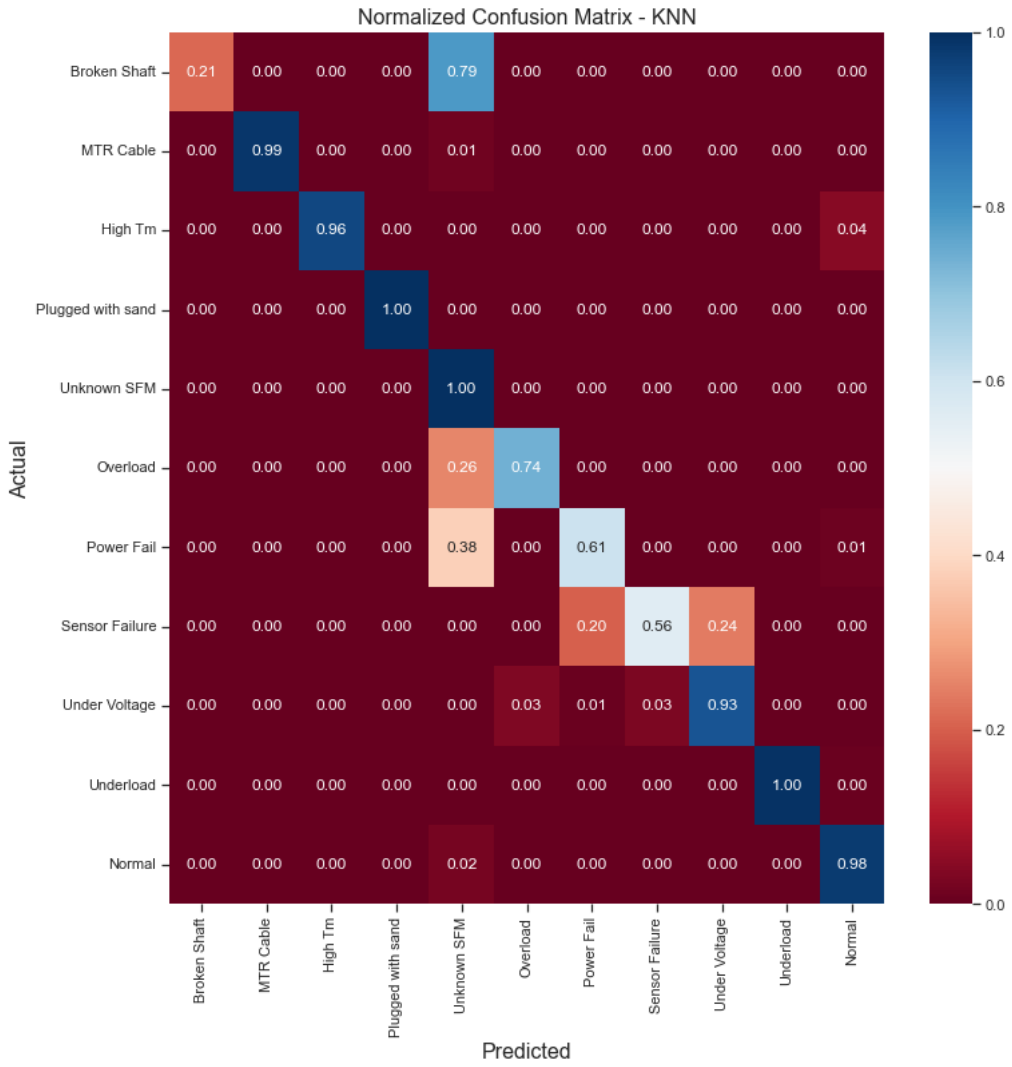

Figure 6-17: Design C-KNN Normalized Confusion Matrix (3 Days PP)

Comparing Design C to Design A, it can be deduced that the integration of GVF and pump efficiency has a beneficial effect on prediction of two ESP SFMs, namely motor underload, and high motor temperature. Gas locking is one of the potential causes of motor underload failure. Upon entrance of free gas into the pump, the low-density gas separates from the higher-density fluid, accumulates, and eventually blocks the passage of the whole vane in the pump stage, resulting in gas locking, as explained in Section 2.4.1. Consequently, the liquid flow ceases, resulting in a decreased load on the motor, ultimately resulting in ESP motor underload SFM.

High motor temperature SFM, on the other hand, occurs when free gas replaces the liquid that usually flows around the motor to cool it. This causes a rapid rise in the motor winding temperature and may eventually burn the motor out. The predictions of this failure and motor underload are significantly impacted by gas interference in the ESP system and pump efficiency. Therefore, the incorporation of GVF and pump efficiency as input variables enhances the performance of the ESP SFM prediction.

According to the variable importance presented in Figure 6-18, the motor current, motor frequency, and wellhead pressure have the greatest effect on predicting failures. This result is consistent with the findings of Designs A and B. All the investigated wells in this research are equipped with downhole gas separators, minimizing the effect of GVF on ESP SFM predictions.

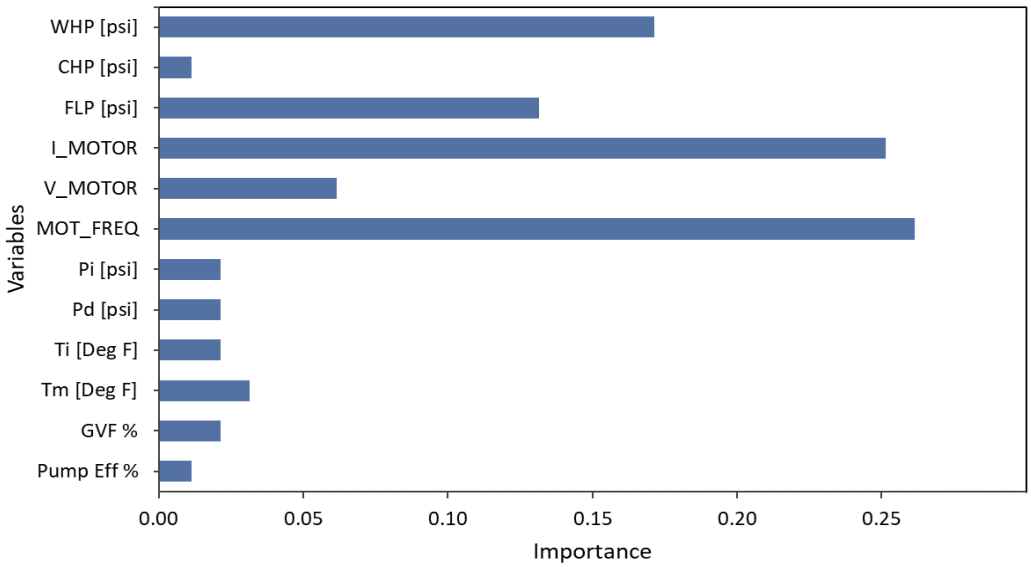

Figure 6-18: Design C-Variables Importance

6.4 Design D

Design D, like Design B, seeks to include physics into the analytical approach and lower the number of variables. It employs input variables from Design B in addition to the pump efficiency, GVF, and input electrical power to serve as the classification model's input variables.

Figure 6-19 summarizes a comparison of average F1-scores of the tested ML models for Design D. K-Nearest Neighbor (K=10) is the best model with the greatest F1-score for predicting ESP Failures throughout all prediction periods and SFM's. Table C- 3 details the tuned parameters for each ML model in Design D. The K value was chosen based on the outcomes of Figure B- 20, Figure B- 21, and Figure B- 22 with K=10 providing the highest precision, recall, and F1-score.

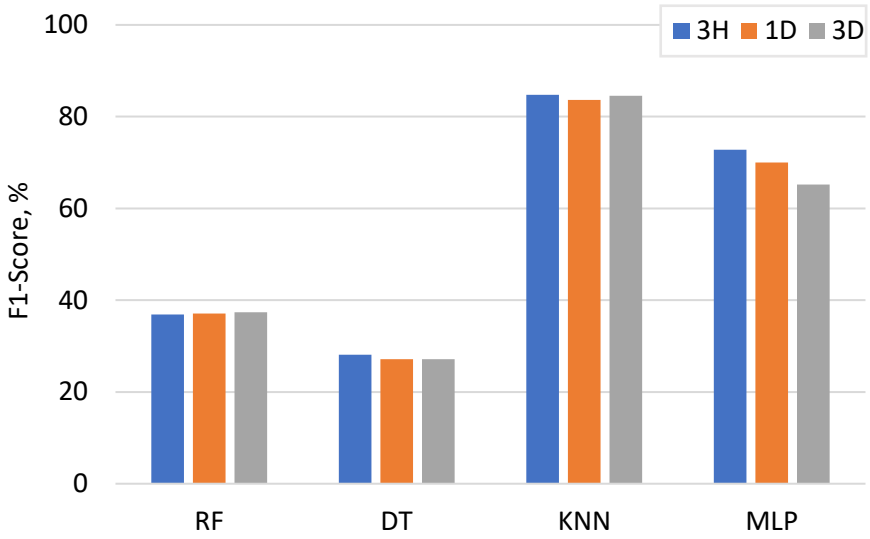

Figure 6-19: Design D- Summary of Model Performance (F1-Score)

Figure 6-20 summarizes a comparison of Cohen's Kappa Coefficient (κ) results of the tested ML models for Design D. The KNN model achieves the highest Kappa (κ) among the other models, which corresponds to an almost perfect agreement between the actual and predicted SFMs (> 0.86).

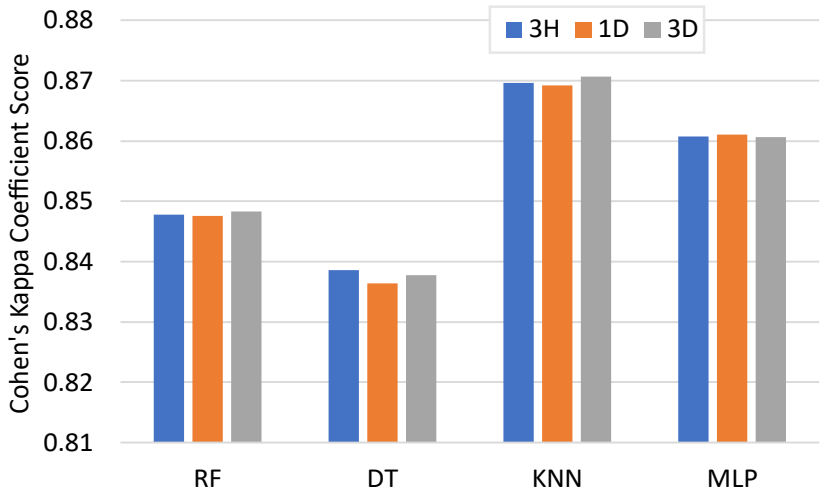

Figure 6-20: Design D- Summary of Cohen's Kappa Coefficient (κ) Results

The correlation heatmap in Figure 6-21 demonstrates that pump efficiency has a strong positive correlation with ΔP_{pump} and motor frequency. This is consistent with the fact that the pump efficiency is a function of pressure difference in the pump (ΔP_{pump}). However, there is a negative relation between the pump efficiency and input electrical power, GVF, and ϕ. As stated before, the greater the GVF, the lower the pump's efficiency becomes owing to gas interference. In addition, the GVF is positively correlated with ϕ, input electrical power, and motor frequency.

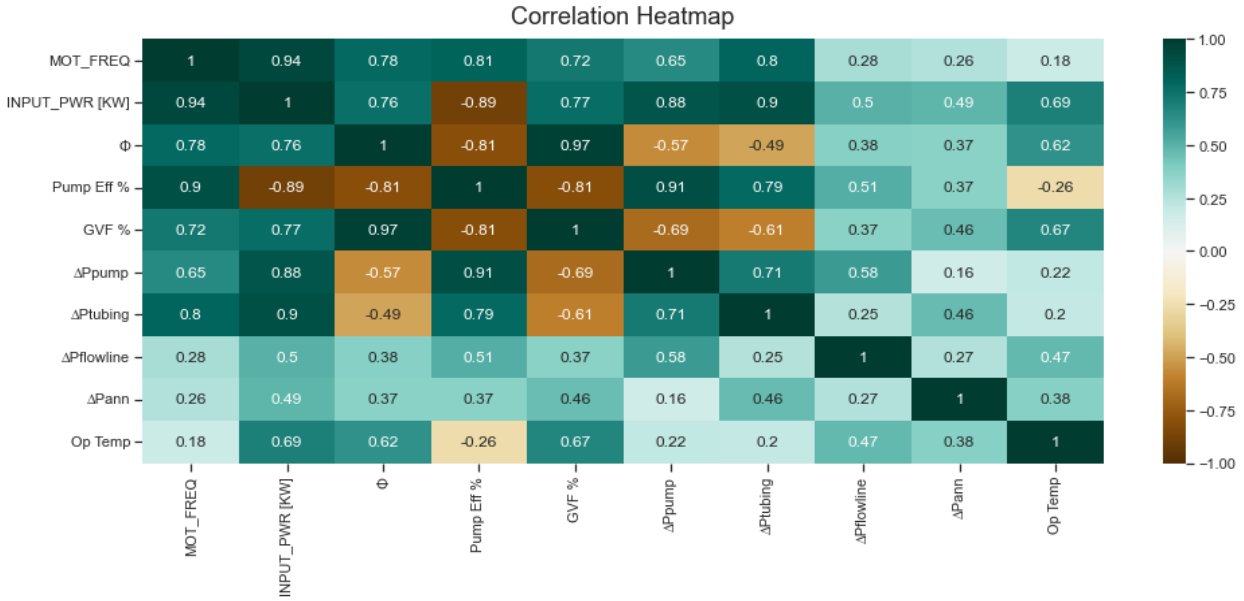

Figure 6-21: Design D- ESP Variables Correlation Heatmap

The variable importance in Figure 6-22 demonstrates that motor frequency and $\Delta P_{flowline}$ have the most influence in predicting failures, which is consistent with the variable importance of Design B. This resemblance is a result of the fact that Design D is constructed with the same input variables as Design B, and as well as GVF, η_{pump}, and input electrical power.

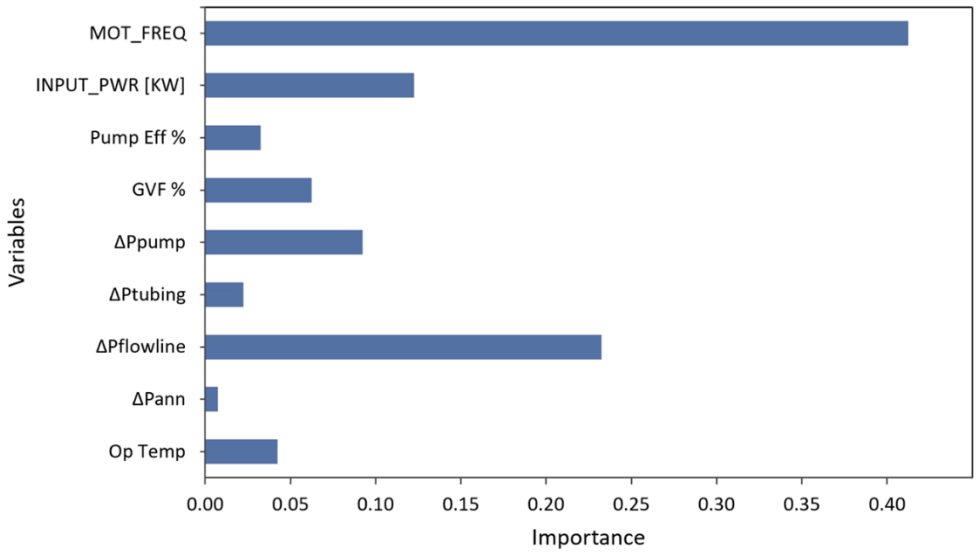

Figure 6-22: Design D-Variables Importance

Table 6-5 describes the performance of Design D's best model, KNN, across the three prediction periods for each ESP specific failure. As observed, the KNN model demonstrates a high degree of precision and recall for most SFM's, indicating that the algorithm is performing effectively.

Table 6-5: Design D-KNN Results over Three Prediction Periods

	Prediction Period of 3 hours			Prediction Period of 1 Day			Prediction Period of 3 Days		
SFM	**Precision**	**Recall**	**f1-score**	**precision**	**recall**	**f1-score**	**precision**	**recall**	**f1-score**
Broken Shaft	1	0.21	0.35	1	0.21	0.35	1	0.21	0.35
MTR Cable	0.97	0.9	0.94	0.97	0.9	0.94	0.97	0.97	0.97
High Tm	1	0.94	0.97	0.95	0.86	0.9	0.99	0.97	0.98
Plugged with sand	1	0.99	0.99	0.99	0.98	0.99	0.99	0.99	0.99
Unknown SFM	0.77	1	0.87	0.77	1	0.87	0.77	1	0.87
Overload	0.97	0.73	0.83	0.97	0.73	0.83	0.97	0.73	0.83
Power Fail	0.96	0.59	0.73	0.97	0.6	0.74	0.97	0.6	0.74
Sensor Failure	0.84	0.64	0.73	0.82	0.56	0.67	0.82	0.56	0.67
Under Voltage	0.91	0.93	0.92	0.91	0.92	0.92	0.89	0.92	0.91
Underload	1	1	1	1	0.99	1	1	0.99	1
Normal	1	0.98	0.99	1	0.98	0.99	1	0.98	0.99
Cohen's Kappa Coefficient Score=	0.8696			0.8692			0.8707		

The KNN normalized confusion matrix in Figure 6-23 shows the performance of Design D with a three-day prediction period. Figure B- 24 and Figure B- 23 show similar plots with the three-hour and one-day prediction periods, respectively. The matrices indicate that the model predicts each ESP failure accurately, while exhibiting a low percentage of false alarms. Each diagonal value shows the rate of true positives the model has accurately identified based on its recall measure. Similar to the other designs, broken shaft is the only failure poorly predicted by the model due to the lack of vibration data in this study.

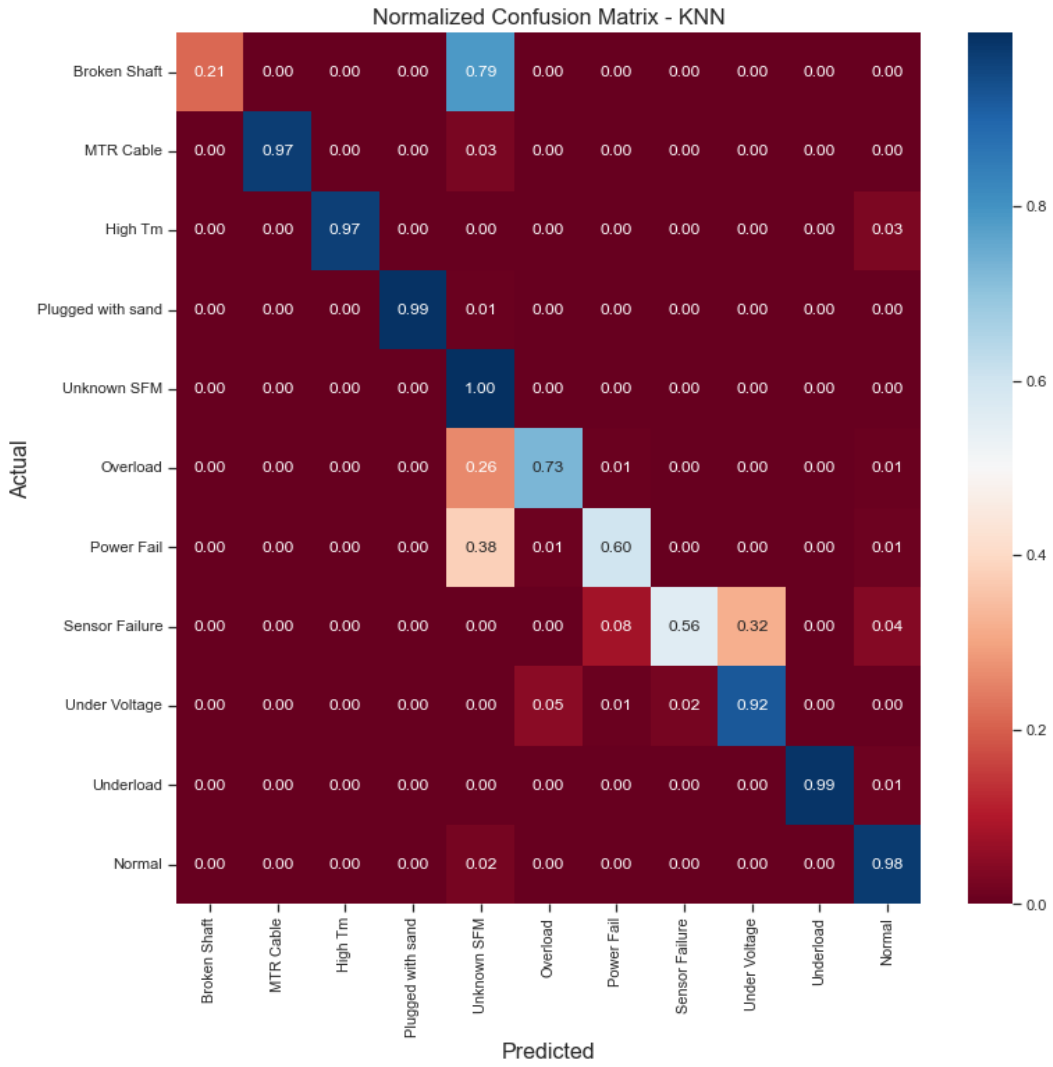

Figure 6-23: Design D-KNN Normalized Confusion Matrix (3 Days PP)

6.5 Design E

The main objective of Design E is to forecast the pump efficiency (η_{pump}) by using a supervised regression model, with ESP-labeled datasets serving as the input variables. Table 6-6 provides a summary of the regression results for the tested ML models of Design E. Multilayer Perceptron (MLP) Neural Network is the best model to forecast the η_{pump} in Design E, shown by its high R^2 and low MAE and RMSE values. Table C- 2 details the tuned parameters for each ML model in Design E. Figure 6-24 compares the predicted and actual pump efficiency within the

testing dataset using the MLP model. The predicted values are quite close to the fitted line, indicating a strong fit for the MLP model with a high R^2 of 99.6%.

Table 6-6: Design E- Summary of Model Performance

#	Model	RMSE	MAE	R^2	EV
1	**KNN**	2.351	0.724	0.973	0.973
2	**SVR**	2.527	0.714	0.969	0.969
3	**RF**	2.445	1.136	0.971	0.971
4	**MLP**	0.873	0.181	0.996	0.996
5	**LASSO**	4.472	2.875	0.904	0.904
6	**Ridge**	4.468	2.870	0.904	0.904
7	**Polynomial**	2.372	0.784	0.973	0.973
8	**Decision Tree**	2.417	0.872	0.972	0.972

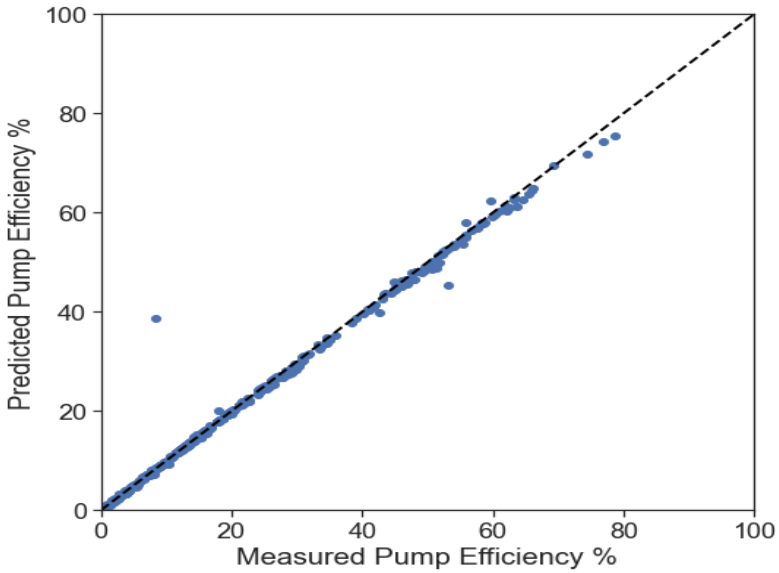

Figure 6-24: MLP-Predicted vs. Measured Pump Efficiency

A multilayer perceptron (MLP) is a kind of feedforward artificial neural network that generates outputs in response to a set of inputs (Pedregosa *et al.*, 2011). In contrast to other machine learning algorithms, MLP allows the user to choose the number of neurons, the activation function, the optimizer, and the learning rate. A well-chosen collection of hyperparameters has the potential to make or break a model's ability to predict the desired outcome accurately. Tuning was performed for the MLP parameters' hyperparameter tuning. The MLP was tuned using the

hyperparameter optimization (GridSearchCV) results and then applied to the ESP dataset as summarized in Table 6-7.

Table 6-7: MLP Hyperparameter Optimization

Hidden Layer Sizes	Activation Function	Alpha	Learning Rate	Mean Test Score	Std Test Score
(50, 50)	tanh	0.0001	Constant	0.996855	0.004382
(50, 50)	tanh	0.0001	Adaptive	0.996855	0.004382
(50, 100)	tanh	0.05	Adaptive	0.996379	0.005773
(50, 100)	tanh	0.05	Constant	0.996379	0.005773
(50, 50)	tanh	0.05	Constant	0.996363	0.004466

Sorting the parameters of Table 6-7 by their test scores, the highest test score was produced by applying the optimum parameters. These parameters include the activation function of "tanh", alpha value of 0.0001, hidden layer sizes of (50, 50), learning rate of "constant", and the "Adam" solver.

The variable importance is depicted in Figure 6-25 for Design E, where the motor voltage (V_m) has the greatest influence on forecasting η_{pump}, followed by ΔP_{pump} as the second most important variable and the oil production rate (Q_o). This result agrees with the concept that the pump efficiency is a function of Q_o, V_m, ΔP_{pump}, as shown in Equation 5-5.

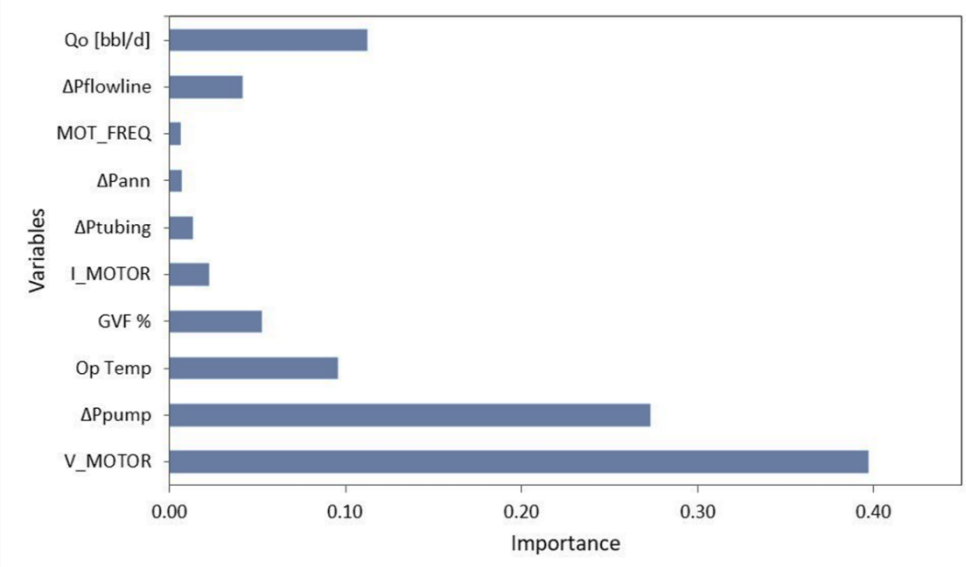

Figure 6-25: Design E-Variables Importance

6.6 Data Analytics Summary

For three days forecast period, Figure 6-26, Figure 6-27, Figure 6-28, and Figure 6-29 highlight the best model results (KNN) for each design covered in this chapter. All designs attain Kappa (κ) values higher than 0.86 which corresponds to an almost perfect agreement between the actual and predicted SFMs, as seen in Figure 6-29. When it comes to accurately forecasting ESP specific failures with the most significant true positive rate, the KNNs models in Designs C and D outperform the models in Design A and B by a small margin. This can be seen in terms of overall F1-score and Kappa coefficient score. Design D is the recommended option for general use since it incorporates the impacts of the effects of gas presence and pump efficiency while using fewer variables than Design C. Although Design D is more favourable, the overall outcomes for all designs are relatively close. We believe Designs C and D have the potential to become much more advantageous if the ML models are applied for high GVF wells without a downhole separator.

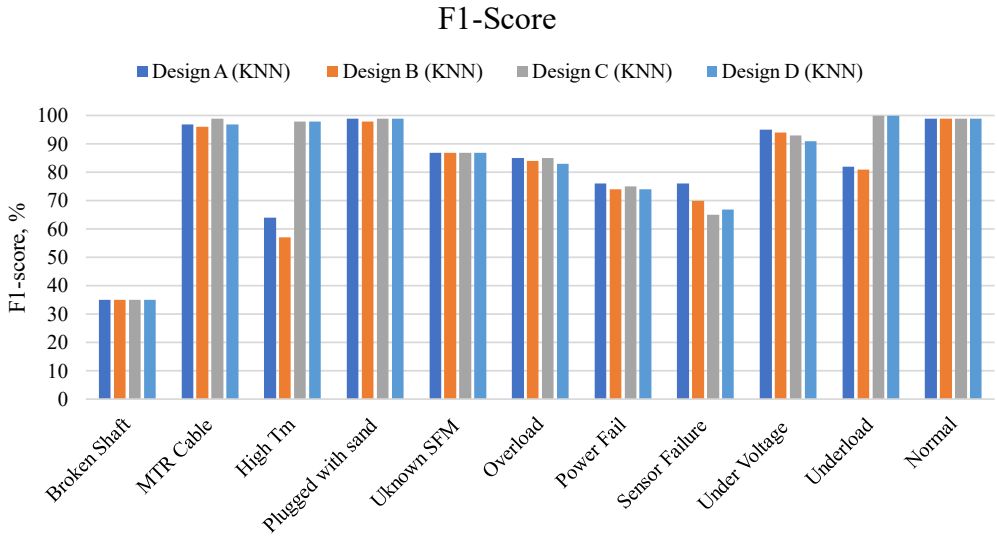

Figure 6-26: F1-Score Results Comparison Across All SFMs for All Designs (3 Days PP)

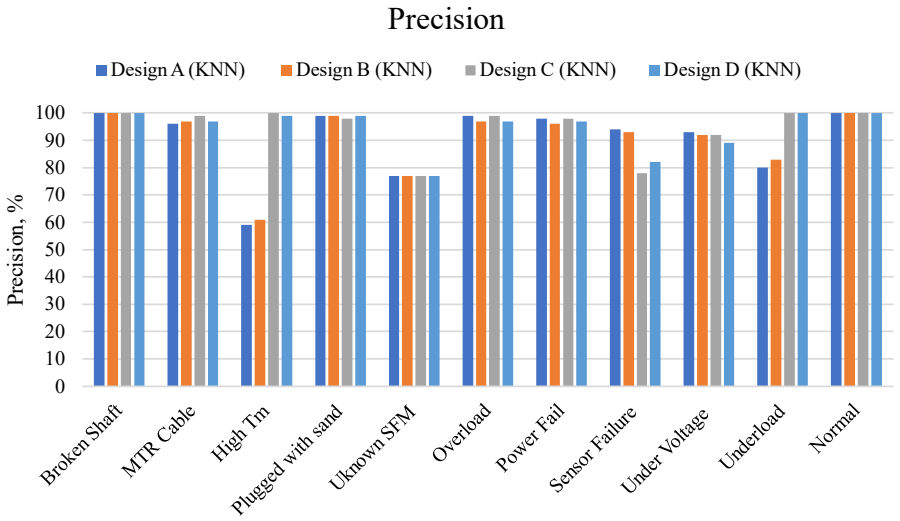

Figure 6-27: Precision Results Comparison Across All SFMs for All Designs (3 Days PP)

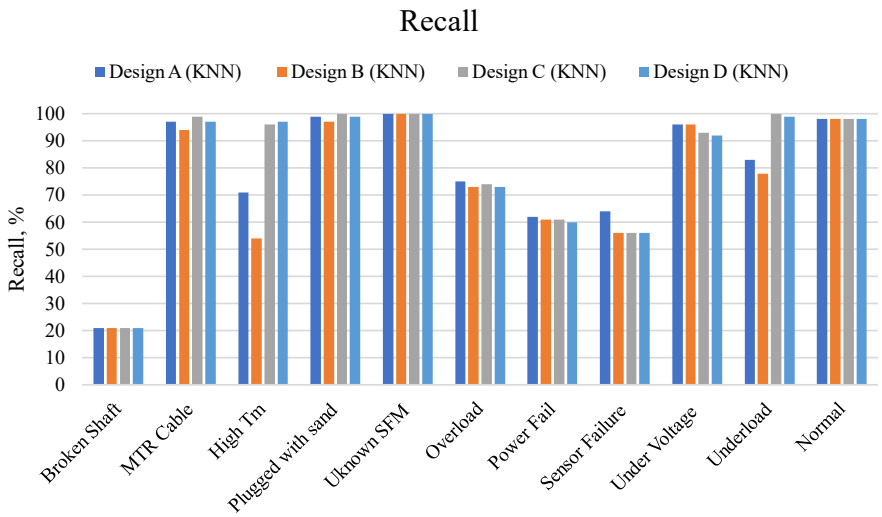

Figure 6-28: Recall Results Comparison Across All SFMs for All Designs (3 Days PP)

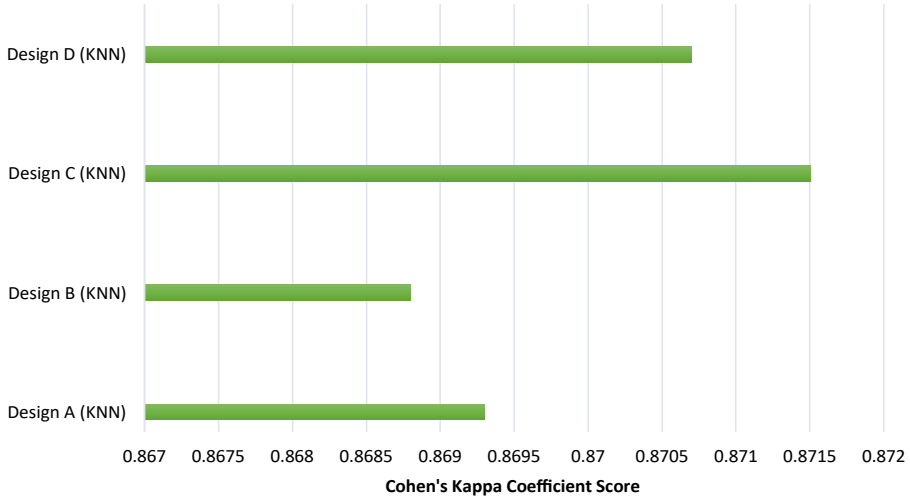

Figure 6-29: Cohen's Kappa Coefficient Score for All Designs (3 Days PP)

Figure 6-30 shows the results of Design D, which is the preferred design for general use in supervised multi-class classification to predict ESP SFMs. The top three ML models (KNN, MLP, and RF) and their F1-score outcomes are compared across each ESP SFM for a three-day prediction period. As seen from Figure 6-30, KNN outperforms MLP and RF across all SFMs despite its simplicity. However, if applied for real-time (streaming) data, KNN may not work efficiently due to its time-consuming and computationally more expensive nature to compute distances and find the nearest neighbors. The second-best model is MLP, which accurately predicts most of the ESP SFMs. MLP handles huge datasets and complex nonlinear problems effectively, but its hyperparameter tuning is challenging, because of the vast number of parameters involved. As shown in Figure 6-30, RF has a lower prediction rate for all SFMs than KNN and MLP, particularly for High T_m and Sensor Failure. Therefore, despite the fact that RF is faster than KNN and MLP and can handle real-time data, it is unable to predict most SFMs with a high F1-score. It should be mentioned that these ML models must be expanded to incorporate streaming data in future applications.

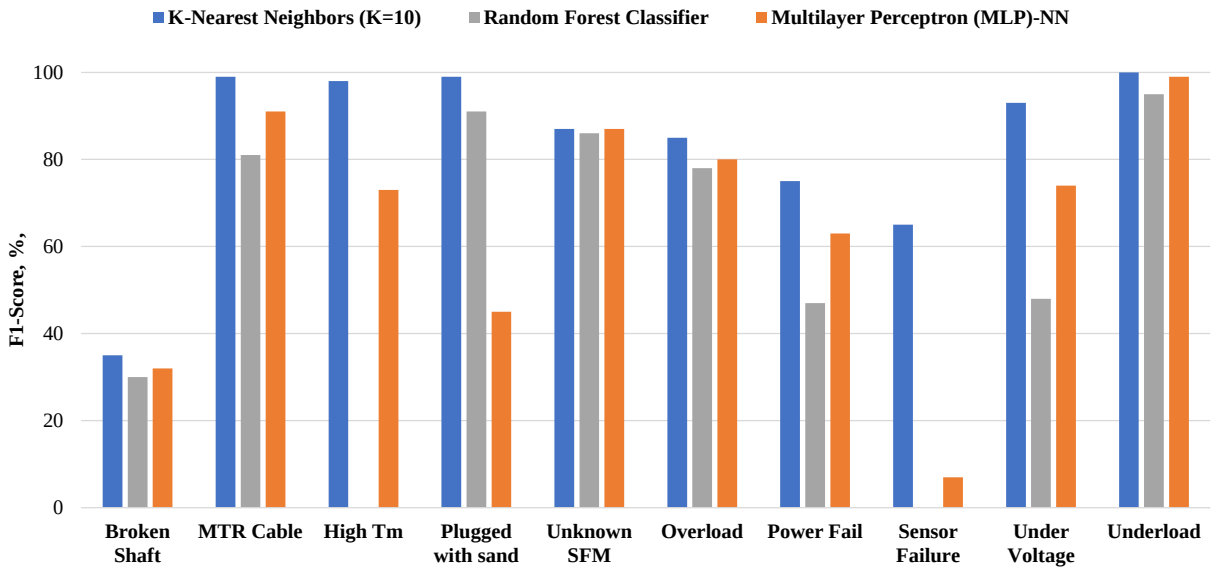

Figure 6-30: Design D-ML Models Comparison (3 days PP)

Chapter 7 Future Applications

As shown in Chapter 6, the four machine learning designs were able to accurately predict ESP specific failures up to three days in advance using the offline dynamic and historical ESP data. The results showed high true prediction rates with a low rate of false alarms. The future challenge will be using these models on live data, re-evaluating and improving their performances, and obtaining daily predictions on active wells. Real-time data processing requires an appropriate selection of ML methods, such as RF, XGBoost, or even deep neural networks. When additional wells are added, this will be a practical execution of this work on a broader scale.

In the real-world implementation of the approach presented in this study, a comprehensive operational database of the field and a record of all ESP failures are necessary. The availability of reliable, organized operational data improves the accuracy of ML model predictions. Hence, the greater the amount of high-resolution data, the better the prediction will be. Many oil and gas companies use their own or third-party data storage and management platforms to handle large amounts of data (Qing *et al*., 2021), using private servers.

Real-time data collected by Supervisory Control and Data Acquisition (SCADA) systems are the keystone of the live machine learning models. The SCADA systems are computer-based systems that control industrial operations by collecting real-time data from distant fields to monitor and control equipment and conditions (Maseda *et al*., 2021). In a SCADA system, the remote terminal units (RTUs) collect data from sensors, counters, meters, and other devices in the field, convert them to digital data, and transmit them to the organization's primary control system

(Maseda *et al*., 2021). After gathering data, the operators may employ analytical tools to get a comprehensive grasp of the data, enabling them to make more informed decisions in less time.

The reliable historical data may be combined with the real-time SCADA data using the existing platforms within the company to provide inputs for the machine learning models. The collected data are then categorized, preprocessed, and prepared for use as training/testing datasets through an online platform. This platform will employ ESP-labeled datasets as inputs and utilize a supervised multi-class classification model to predict several ESP SFM's. These failures are user-defined depending on the field's prevalent ESP failures. A focused team of engineers will evaluate each flagged well for potential failure and verify the alerts by associating them with the probability of ESP failure (Weibull model). Based on the team's findings, proactive ESP maintenance such as optimization, troubleshooting, or pump replacement may be proposed.

Figure 7-1 displays the Live ESP Failure Prediction workflow. This workflow offers an overview of the discussion above on future machine learning model implementations. Shifting from a reactive event-based to a proactive and predictive maintenance of artificial lift operations will aid operators in avoiding undesirable events, significantly reducing downtime, and extending the lifespan of ESP's. ML models are essential in this workflow and cannot be underestimated in our industry's future operations.

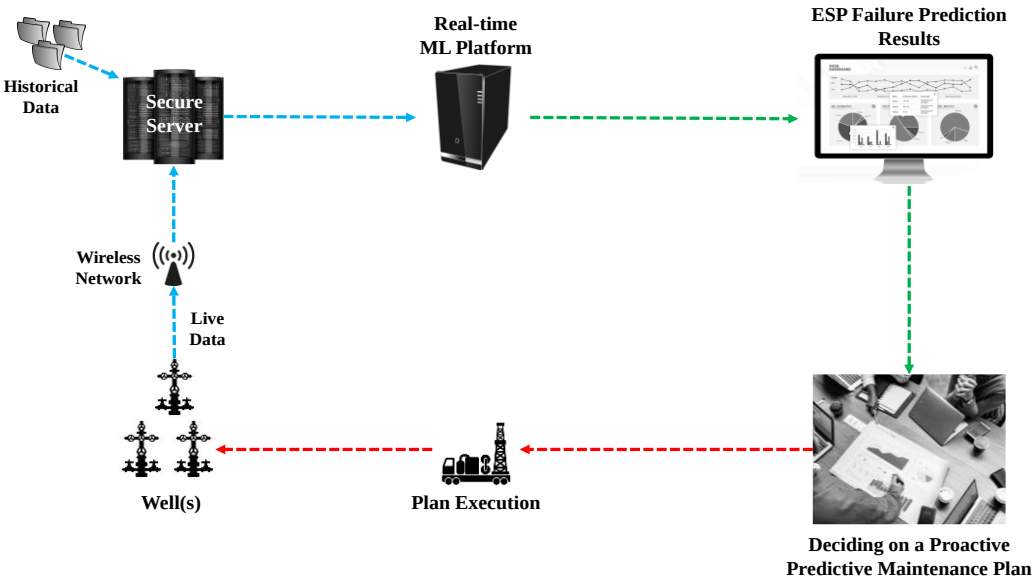

Figure 7-1: Live ESP failure prediction workflow

Chapter 8 Conclusions and Recommendations

This chapter offers a concise overview of the main takeaways from this study. The summary of the study findings and its conclusions are addressed first. Following that is a list of future work recommendations.

8.1 Summary and Conclusions

This study is an attempt to construct a comprehensive machine learning (ML) technique to handle the high frequency of ESP operational data. The model transforms the data into actionable information to predict various ESP failure modes at different prediction periods. Performance metrics were utilized to validate the machine learning models. Four designs with various input parameters were constructed and five prediction periods of 3 hours, 1 day, 3 days, 5 days, and 7 days were evaluated. The results demonstrated valuable outcomes on the application of data analytics in ESP operations. Some of the key outputs for this study are:

- A literature review on the development of ESP failure prediction through time showed the need to use data analytics and accurately predict failures before their occurrences.
- A statistical analysis was performed on the field data obtained, characterizing the general and specific ESP failures. According to these data, three common categories of ESP failures are electrical failures (61%), motor failures (18%), and gas effects (13%). Looking more specifically, power failure, under-voltage, voltage unbalance, and motor underload are the most common occurrences.

- Two case studies from two oil wells in Kuwait Oil Company (KOC) assets were discussed to explore their ESP failures. Both cases share a common general reason for failure but have distinct underlying causes of failure, highlighting the necessity to identify and diagnose the root cause of ESP failure to determine the best mitigation strategy.
- The gas interference in the ESP system and the pump efficiency were examined. An approach for estimating fluid properties from field data was designed to assess gas interference in the pump.
- The early warning signs of failures were investigated by looking at the data trends within the two weeks before each SFM. On average, the warning signals of ESP failures manifest themselves two days before. The most impacted variables are the motor current and voltage, consistent with the fact that most ESP failures are electrical for the data under investigation.
- The KNN model consistently outperformed other ML models in accurately predicting ESP failures for all the prediction periods, with the highest true prediction rate, F1-score, and Kappa coefficient.
- All tested designs showed similarly good performances in predicting ESP failures. Considering that Design D uses less variables than Design C, and it adds the effects of gas presence and pump efficiency to Design B, it is the recommended option for general use.
- Increasing the prediction period resulted in a negligible drop in the model's performance up to 3 days, showing that the model can predict ESP failures accurately 3 days before their occurrences. However, when the prediction time was increased more than 3 days, the forecasts showed increases in the missed failures and false alarms. Consequently, a prediction period of 3 days was chosen for high reliability of failure predictions.

- In Design E, the ESP-labelled datasets were utilized as inputs, and a supervised regression model was employed to forecast the η_{pump}. The best model with the highest accuracy to predict the pump efficiency was Multilayer Perceptron (MLP) Neural Network with an R^2 of 99.6%, MAE of 0.18, and RMSE of 0.873. The motor voltage (V_m) has the greatest influence on forecasting η_{pump}, followed by ΔP_{pump} and the daily oil production rate (Q_o).
- The Weibull statistical analysis was used to predict the probability of ESP failure and the MTBF values. The ESP's MTBF trends for the ten wells were estimated over five years. Initially, all wells had long MTBF values (higher reliability). The MTBF drops with time, starting from the third year of operation. Consequently, ESP loses its reliability over time as operational issues and pump wear cause an increasing number of failures.
- By integrating the outcomes of the ESP Failure prediction ML model with the Weibull unreliability model, a powerful tool is provided. This tool allows the engineers to detect failures early, diagnose potential causes, and propose preventive actions.

8.2 Recommendations

- This study may be extended to incorporate diverse ESP failures from several wells and analyze a broader range of ESP failures. This will become possible by adding more data from wells in different fields.
- Obtaining additional ESP data is advantageous for examining a broader variety of ESP failures. As an example, the vibration data can be helpful in identifying mechanical issues like a broken shaft.
- The study can be extended to include other artificial lift techniques, such as sucker rod and progressive cavity pumps, to recognize their failure types and how to anticipate them.

- This study may be improved by creating a model that provides actionable insights for each failure it identifies. This would allow the system to not only detect problems, but also give a preventative measure for the petroleum engineer.

Nomenclature

B_g	=	Gas Formation Volume Factor, ft^3/SCF
B_o	=	Oil Formation Volume Factor, RB/STB
B_{ob}	=	Formation Volume Factor at bubble point, RB/STB
C_o	=	Oil Compressibility, psi^{-1}
CAGR	=	Compound annual growth rate, %
CFD	=	Computational Fluid Dynamics simulation
DCF_y	=	Discounted cash flow at year y
ESP	=	Electrical Submersible Pump
H_{req}	=	Total Head Required, ft
NCF_y	=	Net Cash Flow at year y
P_b	=	Bubble Point Pressure, psi
P_d	=	Pump Discharge Pressure, psi
P_i	=	Pump Intake Pressure, psi
P_r	=	Reduced Pressure, dimensionless
Q_g	=	Gas Rate, RB/D
Q_{max}	=	Maximum Pump Rate, BPD
Q_o	=	Oil Flow Rate, STB/D
R_1	=	Rotor Radius, m
R_s	=	Solution Gas-Oil-Ratio, SCF/STB
R_{sb}	=	Solution Gas-Oil-Ratio at bubble point pressure, SCF/STB
R_{si}	=	Initial Dissolved Gas-Oil-Ratio, SCF/STB
R_{sr}	=	Reduced Solution Gas-Oil-Ratio, SCF/STB
T_d	=	Pump Discharge Temperature, °F
T_i	=	Pump Intake Temperature, °F
T_m	=	Motor Temperature, °F
T_{res}	=	Reservoir Temperature, °F
V_m	=	Motor Voltage, volts

q_s	=	Gas Volumetric Flow Rate at pump intake, BPD
°API	=	Stock tank Oil Gravity, dimensionless
BPD	=	Bbl. per day
CatBoost	=	Categorical Boosting Algorithm
CHP, P_{csg}	=	Casinghead Pressure, psi
DIFA	=	Dismantle, Inspection and Failure Analysis
DT	=	Decision Tree
EV Test	=	Explained Variance of test dataset
f(t)	=	failure density function
FLP, P_{fl}	=	Flowline Pressure, psi
FN	=	False Negative
FP	=	False Positive
GNB	=	Gaussian Naive Bayesian Classification
GOR	=	Gas Oil Ratio, SCF/STB
GVF	=	Gas Void Fraction, Gas Volumetric Fraction
I, I_m	=	Motor Current, Amps
KNN	=	K-Nearest Neighbor
LASSO	=	Least Absolute Shrinkage and Selection Operator
LightGBM	=	Light Gradient Boosted Machine Algorithm
LR	=	Multinomial Logistic Regression
MAE	=	Mean absolute error
ML	=	Machine Learning
MLE	=	Motor Lead Extension
MLP	=	Multilayer Perceptron Neural Network
MSE	=	Mean-square error
MTBF	=	Mean time between failures, days
NN	=	Neural Network
NPV	=	Net Present Value
PCA	=	Principal Component Analysis

PGOR = Portable Gas Oil Ratio

POOH = Pull Out of Hole

P_{pc} = Pseudocritical Pressure, psi

P_{pr} = Pseudoreduced Pressure, dimensionless

Q = Liquid Volumetric Flow Rate at pump intake, BPD

$R(t)$ = reliability function

R^2 = Coefficient of determination

RF = Random Forest Algorithm

RF = Random Forest

RMSE = Root mean-square error

RTU = Remote terminal unit

SCADA = Supervisory Control and Data Acquisition

SFM = Specific Failure Mode of an ESP

SVM = Support Vector Machine Algorithm

SVR = Support Vector Regression

TBF = Time Before Failure, days

TDH = Total Dynamic Head, ft

TN = True Negative

TP = True Positive

T_{pc} = Pseudocritical Temperature, ^{o}R

T_{pr} = Pseudoreduced Temperature, dimensionless

TVD = True Vertical Depth, ft

VSD = Variable Speed Drive

WHP, P_{wh} = Wellhead Pressure, psi

XGBoost = Extreme Gradient Boosting Algorithm

HHP = Hydraulic Horsepower, hp

K = Numbers of nearest neighbourhoods

V = Impeller Volume, m^3

Greek Symbols:

β	=	Shape parameter
η_{pump}	=	Pump Efficiency, %
λ_c	=	Gas Volumetric Fraction (GVF)
ΔP_{ann}	=	Pressure loss in annulus, psi
$\Delta P_{flowline}$	=	Pressure loss in flowline, psi
ΔP_{pump}, ΔP_p	=	Pressure difference in pump, psi
ΔP_{tubing}	=	Pressure loss in tubing, psi
ϕ	=	Turpin Correlation Parameter, psi^{-1}
γ_g	=	Gas specific gravity, dimensionless
γ_o	=	Oil specific gravity, dimensionless
Δ	=	Finite difference operator
κ	=	Cohen's Kappa Statistic
Ω	=	Angular Speed, rad/s
η	=	Scale parameter
λ	=	The failure rate, %
σ	=	Surface Tension, N/m

References

Abdelaziz, M., Lastra, R., & Xiao, J. J. (2017). ESP Data Analytics: Predicting Failures for Improved Production Performance. SPE Paper No. 188513, Abu Dhabi International Petroleum Exhibition & Conference, 15 November, Abu Dhabi, UAE. doi: https://doi.org/10.2118/188513-MS

Adesanwo, M., Denney, T., Lazarus, S., & Bello, O. (2016). Prescriptive-Based Decision Support System for Online Real-Time Electrical Submersible Pump Operations Management. SPE Paper No. 181013, SPE Intelligent Energy International Conference and Exhibition, 6 September, Aberdeen, Scotland, UK. doi: https://doi.org/10.2118/181013-MS

Ahmed, T. H. (2010). Reservoir Engineering Handbook (4th ed.). Gulf Professional Publishing.

Alhanati, F.J.S., Solanki, S.C., and Zahacy, T.A. (2001). ESP Failures: Can We Talk the Same Language?. SPE Paper No. 148333, SPE Gulf Coast Section Electric Submersible Pump Workshop, 25-27 April, Houston, Texas. doi: https://doi.org/10.2118/148333-MS

Al-Sadah, H. (2014). ESP Data Analysis to Enhance Electrical Submersible Pump Run Life at Saudi Arabian Fields. SPE Paper No. 173703, SPE Middle East Artificial Lift Conference and Exhibition, 26-27 November, Manama, Bahrain. doi: https://doi.org/10.2118/173703-MS

Andrade Marin, A., Busaidy, S., Murad, M., Balushi, I., Riyami, A., Jahwari, S., Ghadani, A., Ferdiansyah, E., Shukaili, G., Amri, F., Kumar, N., Marin, E., Gala, R., Rai, R., Venkatesh, B., Bai, B., Kumar, A., Ang, E., & Jacob, G. (2019). ESP Well and Component Failure Prediction in Advance using Engineered Analytics - A Breakthrough in Minimizing Unscheduled Subsurface Deferments. SPE Paper No. 197806, Abu Dhabi International Petroleum Exhibition & Conference,2 November, Abu Dhabi, UAE. doi: https://doi.org/10.2118/197806-MS

Barrios L, Prado MG. (2009). Modeling two phase flow inside an electrical submersible pump stage. Paper OMAE2009-79727 presented at the ASME 28th International Conference on ocean, offshore and Arctic Engineering held in Honolulu; May 31- June 5.

Bermudez, F., Carvajal, G. A., Moricca, G., Dhar, J., Md Adam, F., Al-Jasmi, A., Goel, H. K., & Nasr, H. (2014). Fuzzy Logic Application to Monitor and Predict Unexpected Behavior in Electric Submersible Pumps (Part of the KwIDF Project). SPE Paper No. 167820, SPE Intelligent Energy Conference & Exhibition, 1 April, Utrecht, The Netherlands. doi: https://doi.org/10.2118/167820-MS

Betonico, G., Bannwart, A., M.M., Ganzarolli (2015). Determination of the temperature

distribution of ESP Motors under variable Conditions of flow rate and loading. Journal of Petroleum Science and Engineering. Volume 129, Pages 110-120, ISSN 0920-4105, https://doi.org/10.1016/j.petrol.2015.02.032.

Bhardwaj, A. S., Saraf, R., Nair, G. G., & Vallabhaneni, S. (2019). Real-Time Monitoring and Predictive Failure Identification for Electrical Submersible Pumps. SPE Paper No. 197911, Abu Dhabi International Petroleum Exhibition & Conference, 12 November, Abu Dhabi, UAE. doi: https://doi.org/10.2118/197911-MS

Bravo, C., Rodriguez, J., Saputelli, L., & Rivas, F. (2014). Applying Analytics to Production Workflows: Transforming Integrated Operations into Intelligent Operations. SPE Paper No. 167823, SPE Intelligent Energy Conference & Exhibition, 1 April, Utrecht, The Netherlands. doi: https://doi.org/10.2118/167823-MS

Brill & Beggs, Two Phase Flow in Pipes, 6th Edition, 1991. Chapter 2.

Brown, G.G. (1948). Natural Gasoline and the Volatile Hydrocarbons. Tulsa, Oklahoma: Natural Gasoline Association of America.

Brulé, M. R. (2013). Big Data in E&P: Real-Time Adaptive Analytics and Data-Flow Architecture. SPE Paper No. 163721, SPE Digital Energy Conference, 5 March, The Woodlands, Texas, USA. https://doi.org/10.2118/163721-ms

Birolini, A. (2018). Reliability Engineering Theory and Practice. Springer.

Carrillo, W. (2013). Prognostics for Oil & Gas Artificial Lift applications, GE Oil and Gas. PHM Conference, New Orleans. https://phmsociety.org/wp-content/uploads/2013/08/HMOG_01_GE.pdf.

Cirilo, R. (1998). Air-water Flow through Electric Submersible Pumps. MS Thesis. The University of Tulsa, Tulsa, Oklahoma.

Dunbar, C.E. (1989). Determination of Proper Type of Gas Separator. Paper presented at the Microcomputer Applications in Artificial Lift Workshop sponsored by the SPE Los Angeles Basin Section, October.

Duran, J., Prado, M. (2003). ESP Stages air-water two-phase performance-modeling and experimental data. SPE J., 13, 196105718.

Donner, D., J. Brady, J. Ortstadt, B. Pilot and B. New (2014). October Edition: A Spotlight on the Artificial Lift Market. Oil & Gas Newsletter.

El Gindy, M., Abdelmotaal, H., Botros, K., Ginawi, I., Sayed, E., & Edris, T. (2015). Monitoring & Surveillance Improve ESP Operation and Reduce Workover Frequency. SPE Paper No. 177926, Abu Dhabi International Petroleum Exhibition and Conference, 10 November, Abu Dhabi, UAE. doi: https://doi.org/10.2118/177926-MS

Fakher, S., Khlaifat, A., Hossain, M. E., & Nameer, H. (2021). Rigorous review of electrical submersible pump failure mechanisms and their mitigation measures. Journal of Petroleum Exploration and Production Technology, 11(10), 3799–3814. https://doi.org/10.1007/s13202-021-01271-6

Grassian, D., Bahatem, M., Scott, T., & Olsen, D. (2017). Application of a Fuzzy Expert System to Analyze and Anticipate ESP Failure Modes. SPE Paper No. 188305, Abu Dhabi International Petroleum Exhibition & Conference, 14 November, Abu Dhabi, UAE. doi: https://doi.org/10.2118/188305-MS

Grandini, M., Bagli, E., & Visani, G. (2020). Metrics for Multi-Class Classification: an Overview. ArXiv, abs/2008.05756.

Guo, D., Raghavendra, C.S., Yao, K.T., Harding, M., Anvar, A., Patel, A. (2015). Data Driven Approach to Failure Prediction for Electric Submersible Pump Systems. SPE Paper No.174062, SPE Western Regional Meeting, 27-30 April, Garden Grove, California. doi: https://doi.org/10.2118/174062-MS

Gupta, S., Saputelli, L., and Nikolaou, M. (2016b). "Applying Big Data Analytics to Detect, Diagnose, and Prevent Impending Failures in Electric Submersible Pumps." Paper presented at the SPE Annual Technical Conference and Exhibition, Dubai, UAE, 26 September. doi: https://doi.org/10.2118/181510-MS

Gupta, S., Saputelli, L., and Nikolaou, M. (2016c). Big Data Analytics Workflow to Safeguard ESP Operations in Real-Time. SPE Paper No. 181224, SPE North America Artificial Lift Conference and Exhibition, 17 October, The Woodlands, Texas, USA. doi: https://doi.org/10.2118/181224-MS

Gupta, S., Nikolaou, M., Saputelli, L., Bravo, C. (2016a). ESP Health Monitoring KPI: A Real-Time Predictive Analytics Application. SPE Paper No. 181009, SPE Intelligent Energy International Conference and Exhibition, 6-8 September, Aberdeen, Scotland. doi: https://doi.org/10.2118/181009-MS

Higgs, G. (1994). ESP Performance a Statistical Review. BP Exploration, 3rd European Electrical Submersible Pump Round Table, Aberdeen.

Hughes, B. (2020). Submersible Pump Handbook.

Igwilo, K. C., Okoro, E. E., & Ubanatu, S. (2018). Comparative approach to optimum selection of artificial lift system. Petroleum & Coal, 60 (3).

Khabibullin, R. A., Shabonas, A. R., Gurbatov, N. S., & Timonov, A. V. (2020). Prediction of ESPs Failure Using ML at Western Siberia Oilfields with Large Number of Wells. SPE Paper No. 201881, SPE Russian Petroleum Technology Conference, 26 October, Virtual. doi: https://doi.org/10.2118/201881-MS

Lea, J. F. (2007). Artificial lift selection. Chapter 10 in SPE petroleum engineering handbook, Vol. IV. Society of Petroleum Engineers, Dallas.

Liu, Y., Yao, K., Liu, S., Raghavendra, C. S., Lenz, T. L., Olabinjo, L., Seren, B., Seddighrad, S., and C. G. Dinesh Babu (2010). Failure Prediction for Rod Pump Artificial Lift Systems. SPE Paper No. 133545, SPE Western Regional Meeting, 27 May, Anaheim, California, USA. doi: https://doi.org/10.2118/133545-MS

Maseda, F. J., López, I., Martija, I., Alkorta, P., Garrido, A. J., & Garrido, I. (2021). Sensors data analysis in supervisory control and data acquisition (SCADA) systems to foresee failures with an undetermined origin. Sensors, 21(8), 2762.

McHugh, M. L. (2012). Interrater reliability: the kappa statistic. Biochemia Medica, 22(3), 276–282. https://doi.org/10.11613/bm.2012.031

Mohrbacher, J. D. (1984). A Field Study of ESP Performance in a Deep, Hot, and Sour Environment. SPE Paper No. 12913, SPE Rocky Mountain Regional Meeting,21 May, Casper, Wyoming. doi: https://doi.org/10.2118/12913-MS

Mordor Intelligence (2022). Artificial lift systems market size, trends: 2022 – 27: Industry growth. Retrieved from https://www.mordorintelligence.com/industry-reports/global-artificial-lift-systems-market-industry

Mian, M.A. (2011). Project economics and decision analysis Volume I: Deterministic models, edition 2. Tulsa, Oklahoma: PennWell.

Mubarak, H. A., Khan, F. A.; Oskay, M. M. (2003). ESP Failures / Analysis / Solutions in Divided Zone - Case Study. SPE Paper No. 81488, Middle East Oil Show, 9 June, Bahrain. doi: https://doi.org/10.2118/81488-MS

Nguyen, T. (2020). Artificial Lift Selection Methodology for Vertical and Horizontal Wells in Conventional and Unconventional Reservoirs. In Artificial Lift Methods (pp. 317-347). Springer, Cham.

Oliveira, L.F., Bardy, M. B., Filho, S. S., Neves, E. A., Silva, J. A., Agustoni, J. A., Amaral, A. P. (1997). Analysis of ESP Failure Data from the Northeastern Pole of the Campos Oil Basin. SPE Electrical Sumbersible Pump Workshop, 30 April - 2 May, Houston, Texas.

Patterson, M. M. (1993). A Model for Estimating the Life of Electrical Submersible Pumps. SPE Paper No. 22790, SPE Production & Facilities, 8(04), 247–250. https://doi.org/10.2118/22790-pa

Pennel, M., Hsiung, J., Putcha, V. B. (2018). Detecting Failures and Optimizing Performance in Artificial Lift Using Machine Learning Models. SPE Paper No. 190090, SPE Western Regional Meeting, 22-25 April, Garden Grove, California, USA. doi: https://doi.org/10.2118/190090-MS

Pedregosa, F. and Varoquaux, G. and Gramfort, A., Michel, V., Thirion, B. and Grisel, O., Blondel, M., Prettenhofer, P. and Weiss, R. and Dubourg, V. and Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E. (2011). Journal of Machine Learning Research 12, pp. 2825-2830.

Qing, P., Li, Y., Luo, S., & Xu, Z. (2021). Big Data Platform for Oil and Gas Production Based on Apache Spark. In Modern Industrial IoT, Big Data and Supply Chain (pp. 129-141). Springer, Singapore.

Sawaryn, S. J. (2003). The Dynamics of Electrical-Submersible-Pump Populations and the Implication for Dual-ESP Systems. SPE Paper No. 87232, SPE Production & Facilities, 18(04), 236–246. https://doi.org/10.2118/87232-pa

Sawaryn, S. J., Grames, K. N., Whelehan, O. P. (2002). The Analysis and Prediction of Electric Submersible Pump Failures in the Milne Point Field, Alaska. SPE Paper No. 74685, SPE Production & Facilities, 17(01), 53–61. https://doi.org/10.2118/74685-pa

Sawaryn, S. J., & Ziegel, E. (2001). Statistical Assessment and Management of Uncertainty in the Number of Electric-Submersible Pump Failures in a Field. SPE Paper No. 71551, SPE Annual Technical Conference and Exhibition, 30 September, New Orleans, Louisiana. doi: https://doi.org/10.2118/71551-MS

Sawaryn, S. J., Norrell, K. S., Whelehan, O. P. (1999). The Analysis and Prediction of Electric-Submersible-Pump Failures in the Milne Point Field. SPE Paper No. 1999, SPE Annual Technical Conference and Exhibition, 3 October, Houston, Texas. doi: https://doi.org/10.2118/56663-MS

Sherif, S., Adenike, O., Obehi, E., Funso, A., & Eyituoyo, B. (2019). Predictive Data Analytics for Effective Electric Submersible Pump Management. SPE Paper No. 198759, SPE Nigeria Annual International Conference and Exhibition, 5-7 August, Lagos, Nigeria. doi: https://doi.org/10.2118/198759-MS

Standing, MB. & Katz, DL. (1942). Density of natural gases. Trans. AIME;146(01):140-9.

Standing, M.B. (1947). A Pressure-Volume-Temperature Correlation for Mixtures of California Oils and Gases; American Petroleum Institute: New York, NY, USA.

Stone, P. (2007). Introducing Predictive Analytics: Opportunities. SPE Paper No. 106865, Digital Energy Conference and Exhibition, 11-12 April, Houston, Texas, U.S.A. doi: https://doi.org/10.2118/106865-ms

Tabe, F.L. (1984). An Overview of the Installation, Operation, Maintenance, and Problems Associated with Electrical Submersible Pump System in Central Sumatra, Indonesia. SPE Paper No. 13201, SPE Annual Technical Conference and Exhibition, 16 September, Houston, Texas. doi: https://doi.org/10.2118/13201-MS

Takacs, G. (2017). Electrical Submersible Pumps Manual: Design, Operations, and Maintenance. United Kingdom: Elsevier Science.

Takacs, G. (2009). Electrical Submersible Pumps Manual: Design, Operations, and Maintenance. Netherlands: Elsevier Science.

Turpin JL, Lea JF, Bearden JL. (1986). Gas-liquid flow through centrifugal pumps correlation of data. Proc. 3rd international pump symposium, May, Texas A&M University.

Vazquez, M.E., and Beggs, H.D. (1980). Correlations for Fluid Physical Property Prediction. JPT (June) 968-70.

Velarde, J. J. (1996). Correlation of black oil properties at pressures below the bubble-point. OAKTrust Home. Retrieved March 23, 2022, from https://oaktrust.library.tamu.edu/handle/1969.1/ETD-TAMU-1996-THESIS-V45

Venkataraman, G. and Mikus, T. (1994). Reliability Analysis of. Electrical Submersible Pumps. Shell Development Company, 1994 SPE ESP Workshop, Houston, Texas.

Williams, A.J., Cudmore, J., Beattie, S. (2003). ESP monitoring – where is your speedometer?. ESP workshop, 30 April – 2 May, Houston, Texas.

Zhu J, Zhang H-Q. (2014). CFD simulation of ESP performance and bubble size estimation under gassy conditions. Paper SPE 170727 presented at the annual technical conference and exhibition held in Amsterdam, The Netherlands; October 27-29.

Zhu, J.; Zhang, H.Q. (2018). A review of experiments and modeling of gas-liquid flow in electrical submersible pumps. Energies,11, 180.

Zhu, J.; Zhang, J.; Cao, G.; Zhao, Q.; Peng, J.; Zhu, H.; Zhang, H.Q. (2019). Modeling flow pattern transitions in electrical submersible pump under gassy flow conditions. J. Pet. Sci. Eng., 180, 471–484.

Appendix A

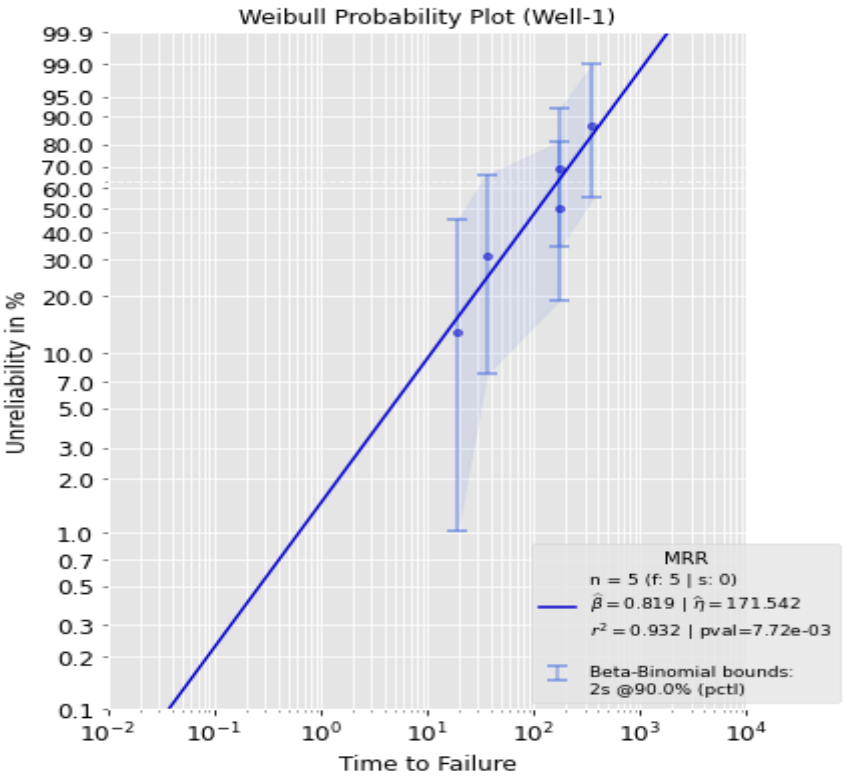

Figure A- 1: Weibull Probability Plot of Well 1

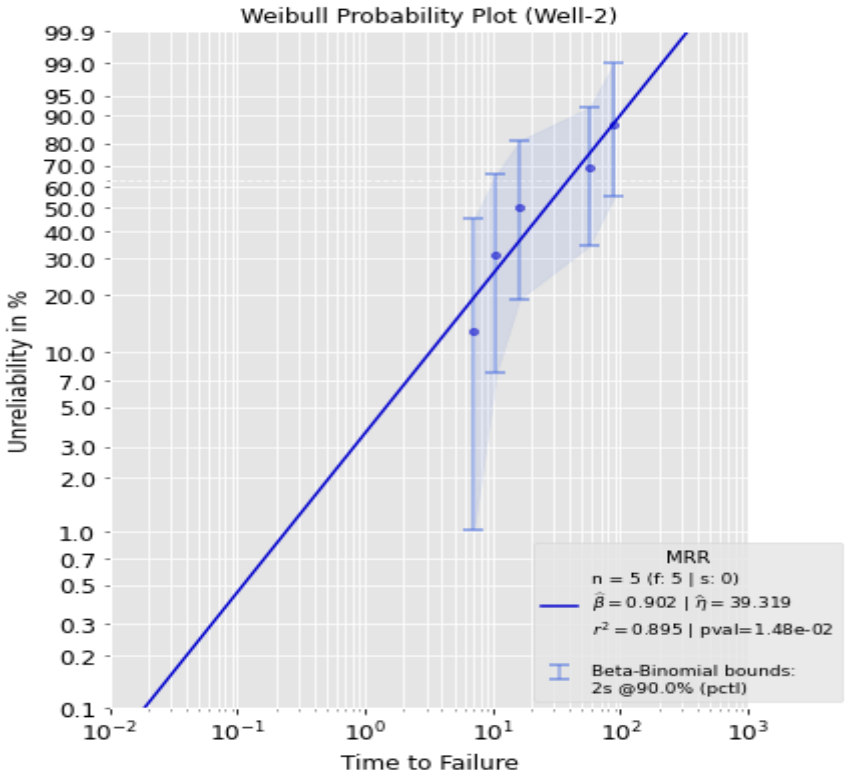

Figure A- 2: Weibull Probability Plot of Well 2

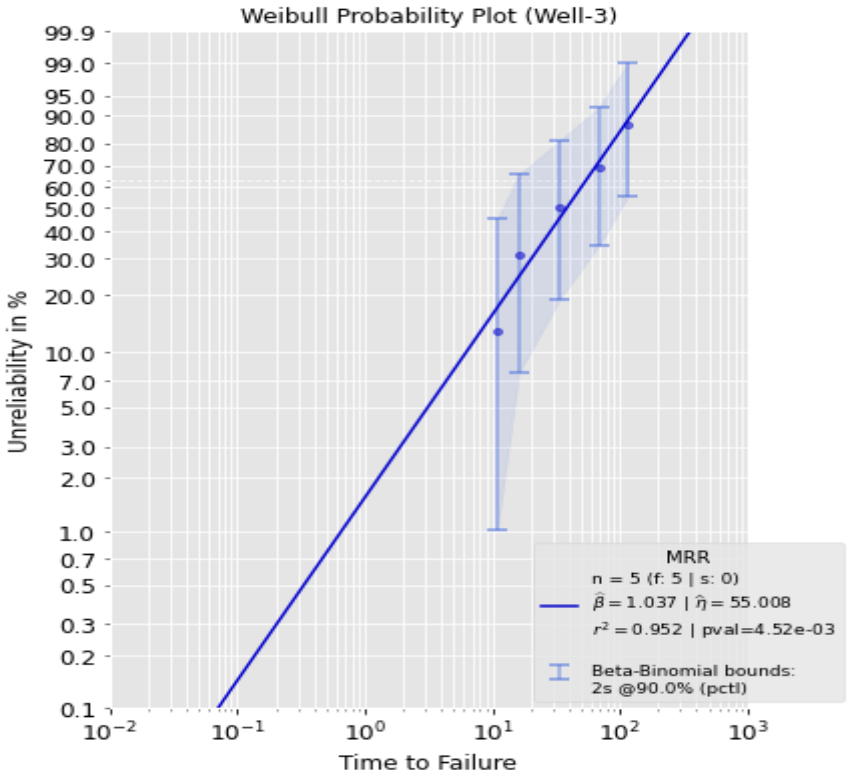

Figure A- 3: Weibull Probability Plot of Well 3

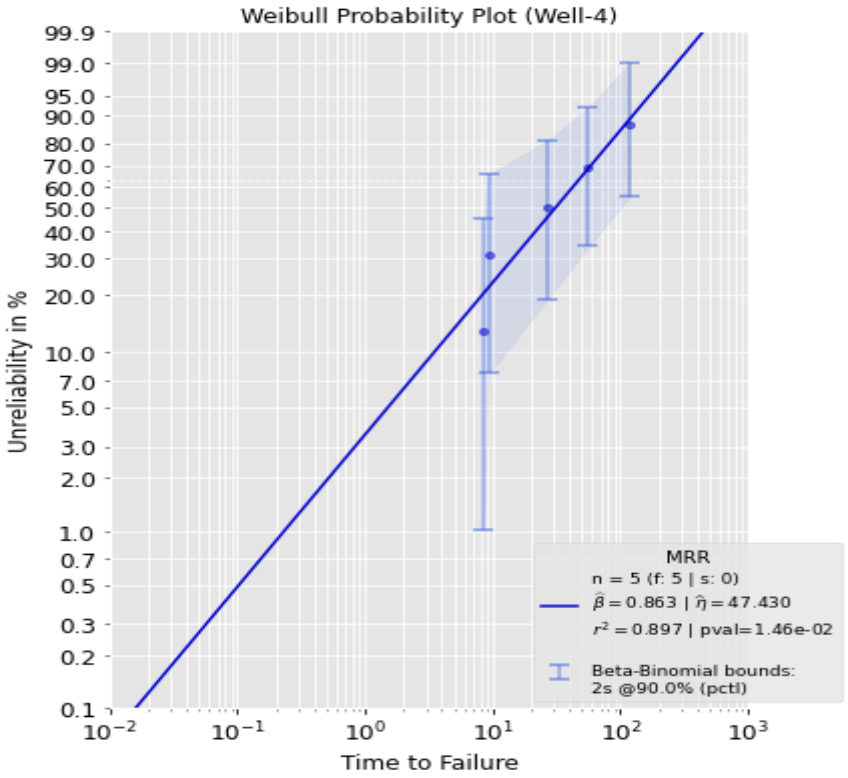

Figure A- 4: Weibull Probability Plot of Well 4

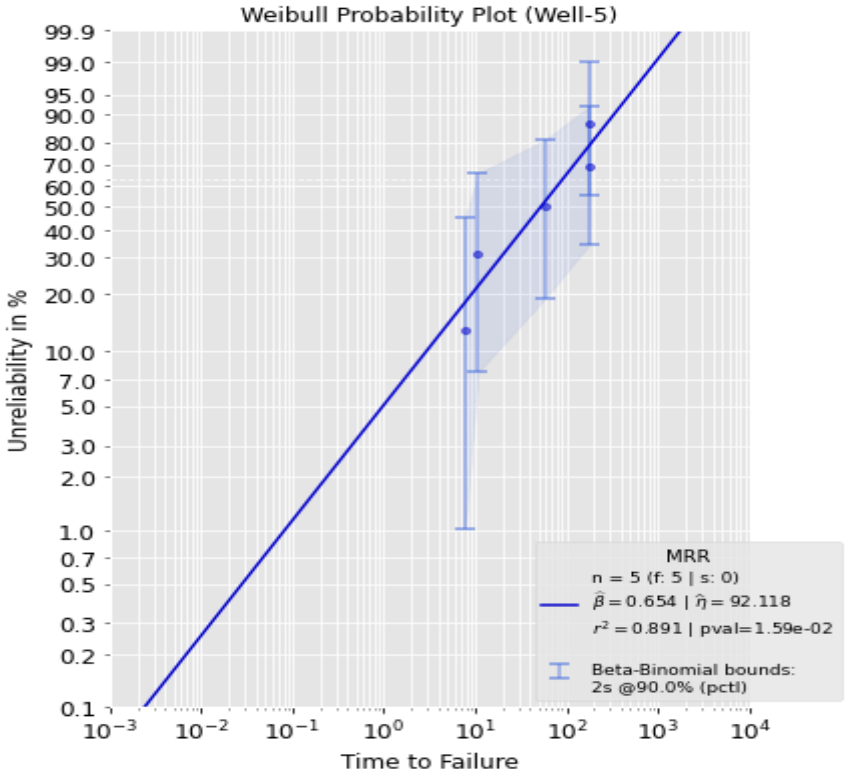

Figure A- 5: Weibull Probability Plot of Well 5

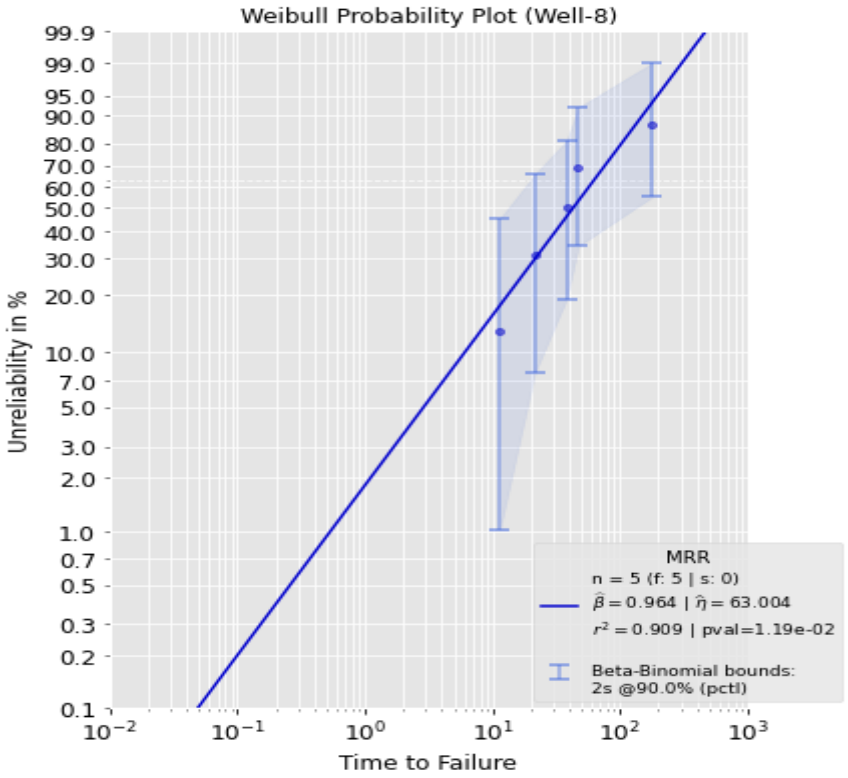

Figure A- 6: Weibull Probability Plot of Well 8

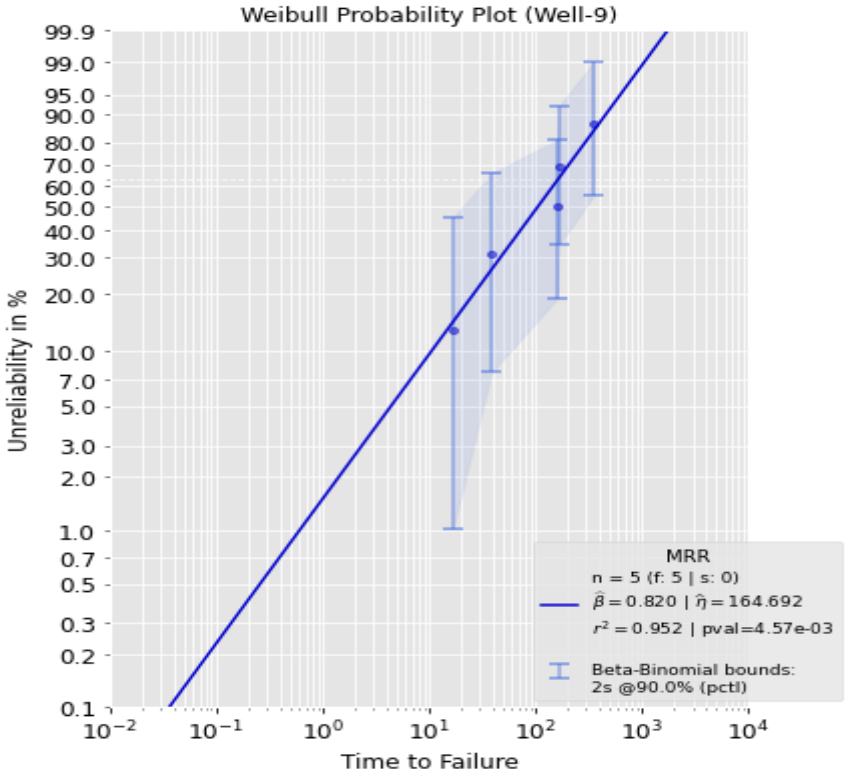

Figure A- 7: Weibull Probability Plot of Well 9

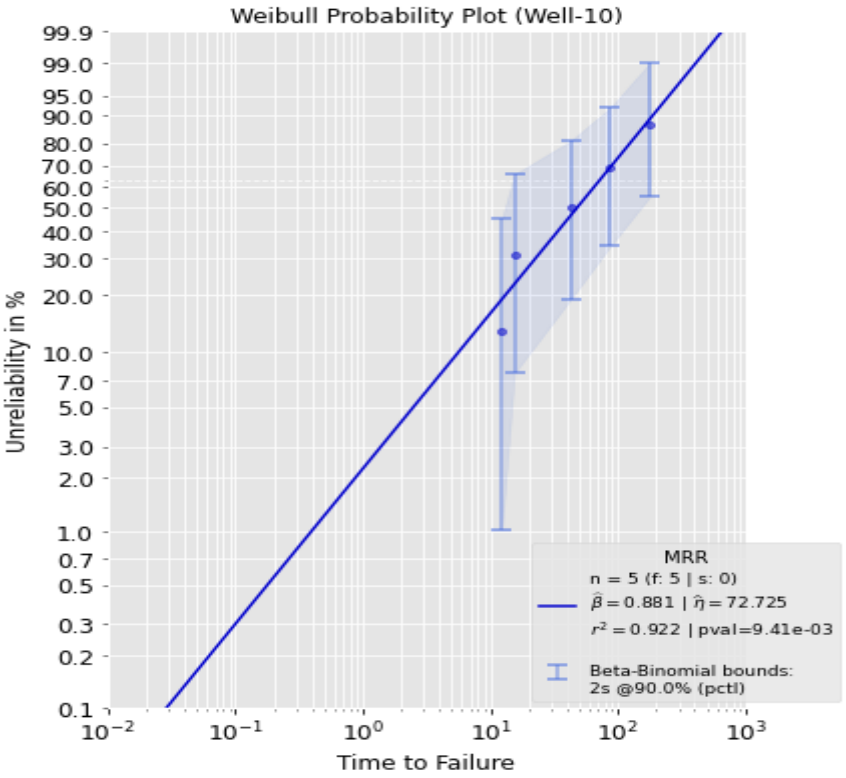

Figure A- 8: Weibull Probability Plot of Well 10

Appendix B

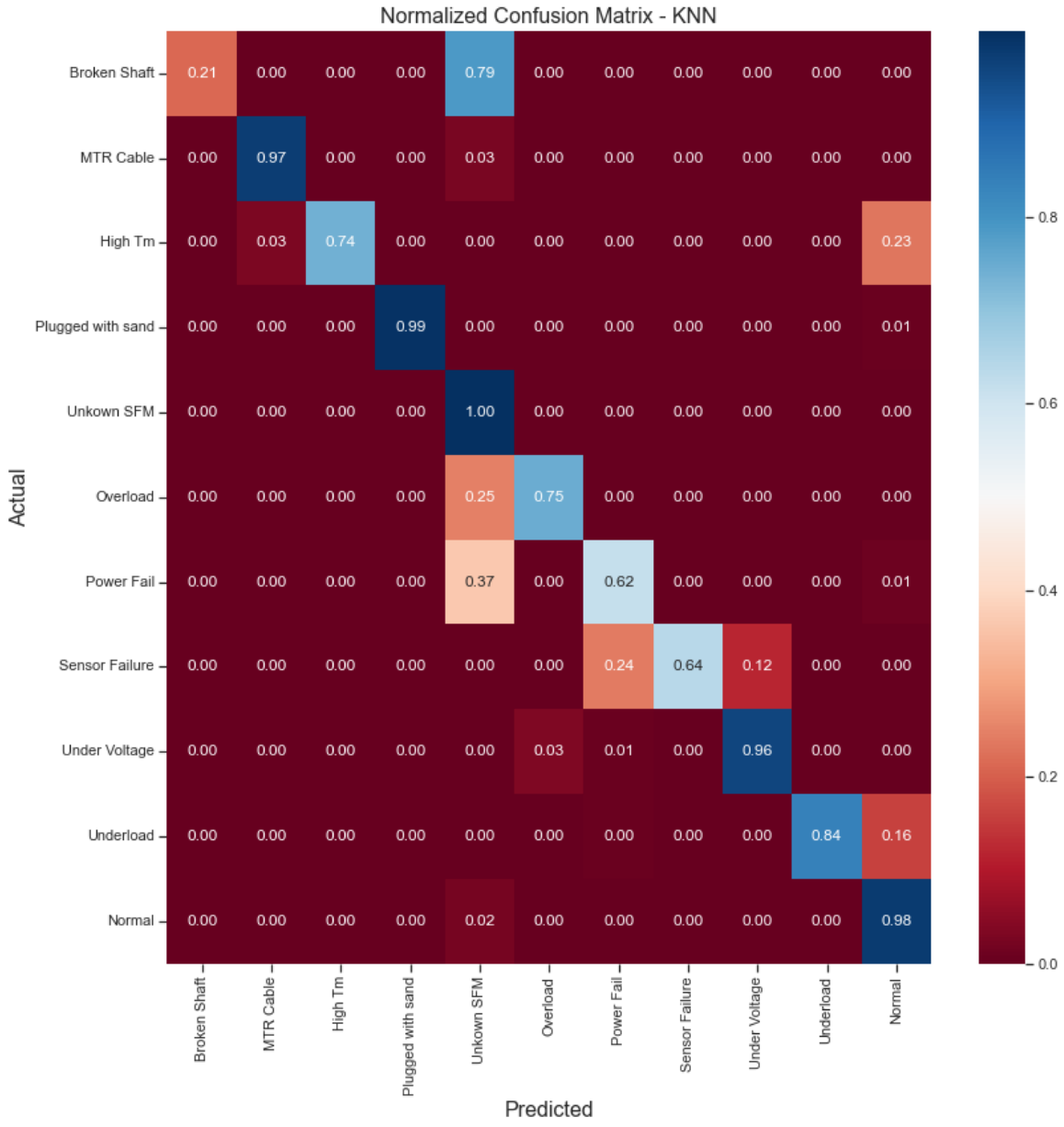

Figure B- 1: Design A-KNN Normalized Confusion Matrix (3 Hours PP)

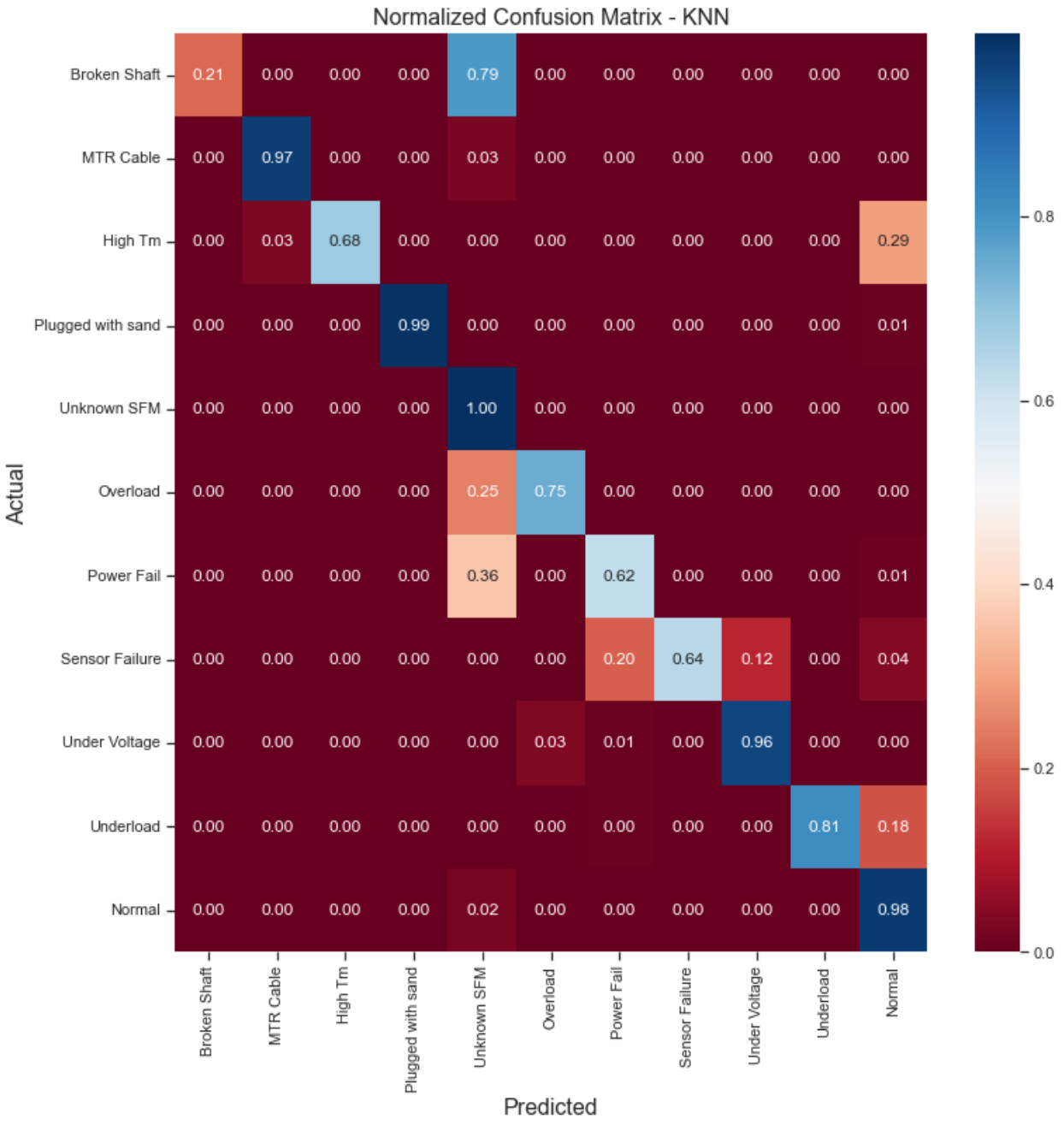

Figure B- 2: Design A-KNN Normalized Confusion Matrix (1 Day PP)

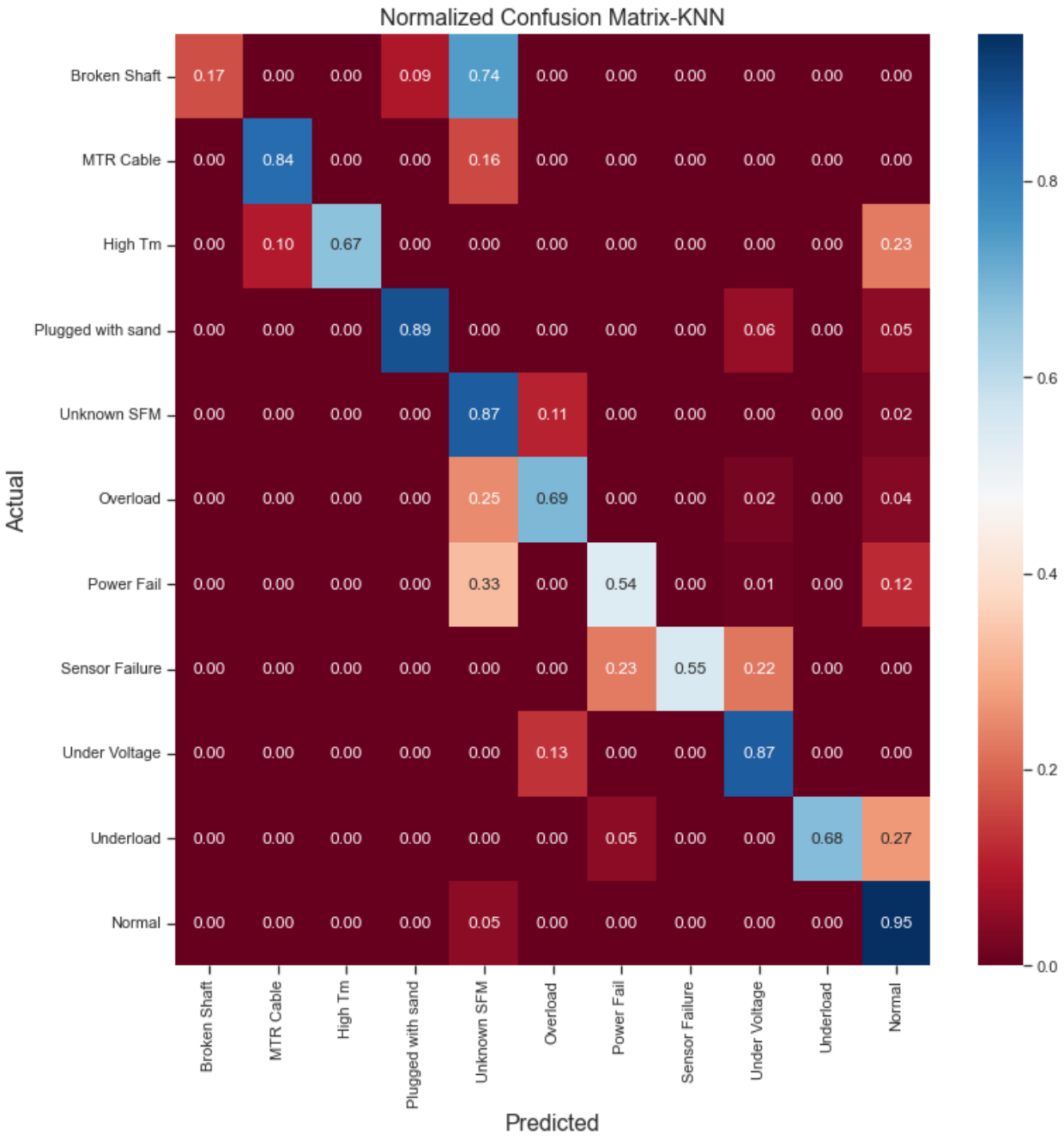

Figure B- 3:Design A-KNN Normalized Confusion Matrix (5-Days PP)

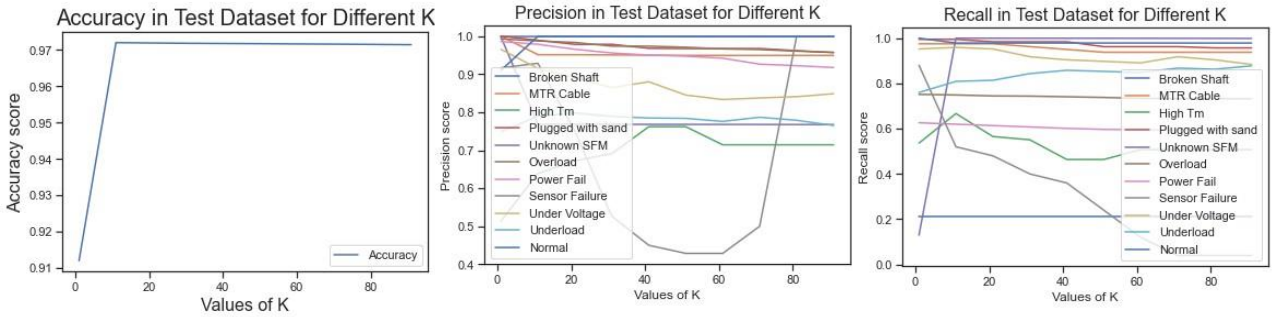

Figure B- 4:Design A-K Values Determination for KNN Model (3-hours PP)

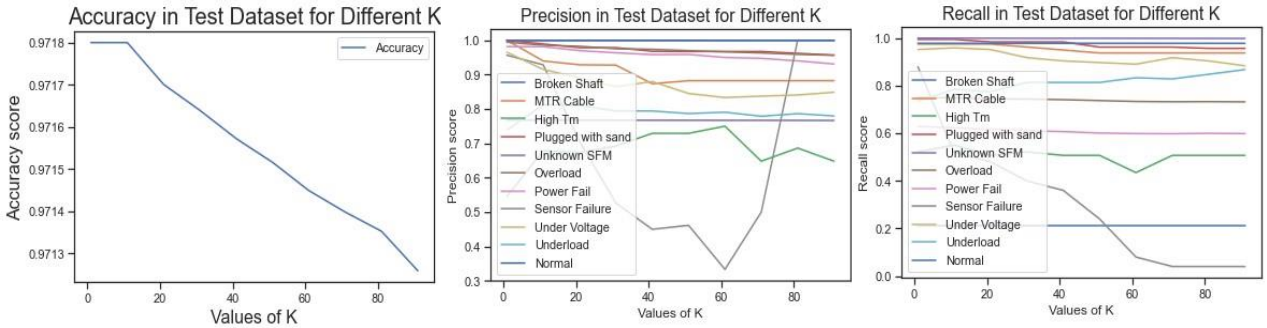

Figure B- 5: Design A-K Values Determination for KNN Model (1 Day PP)

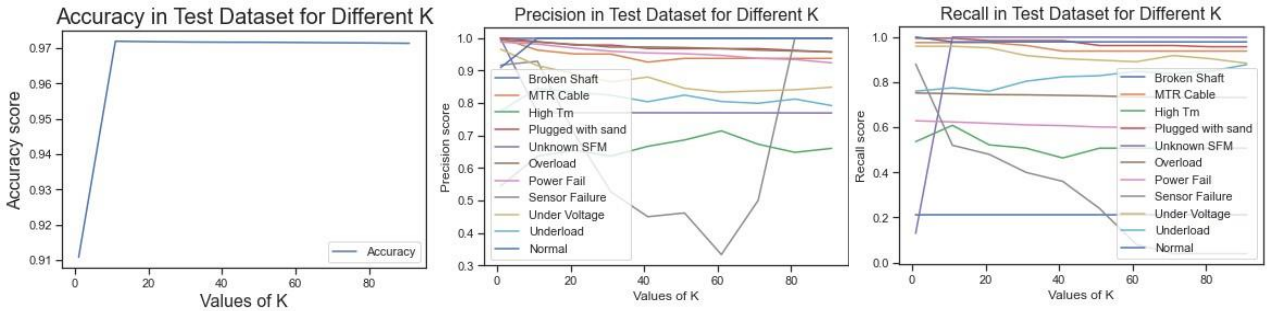

Figure B- 6: Design A-K Values Determination for KNN Model (3 Days PP)

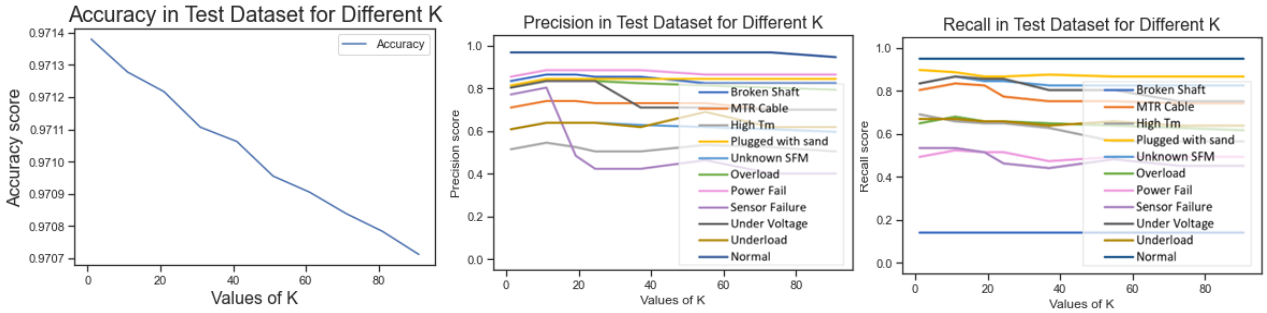

Figure B- 7: Design A-K Values Determination for KNN Model (5 Days PP)

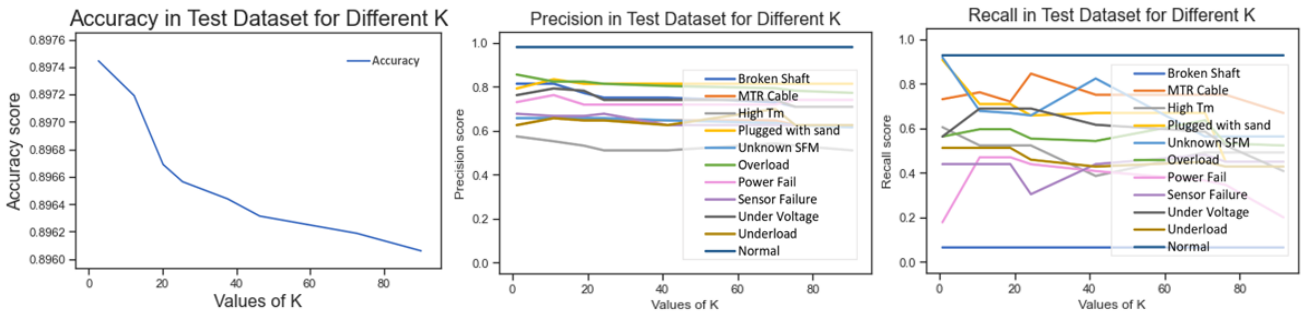

Figure B- 8: Design A-K Values Determination for KNN Model (7 Days PP)

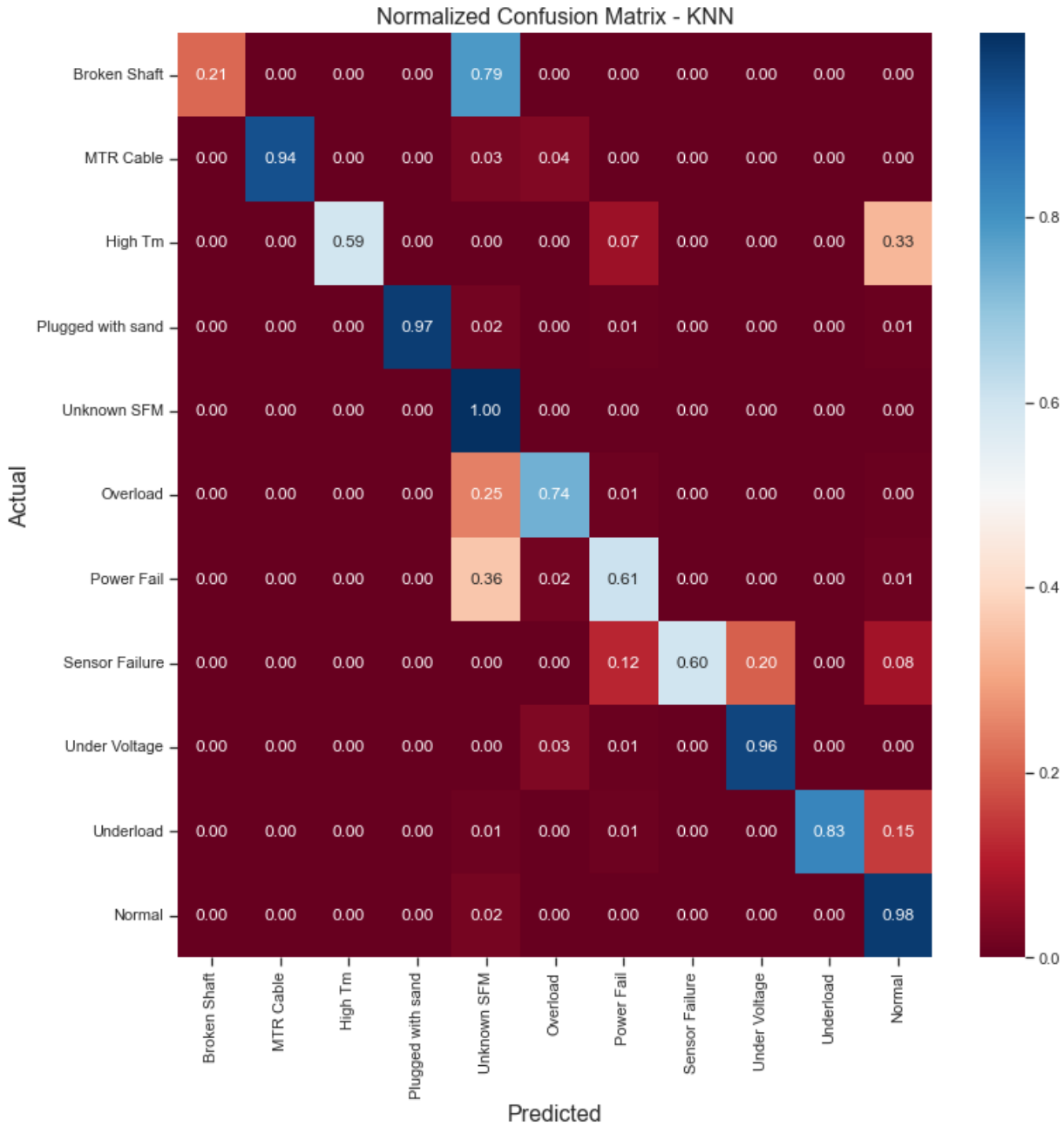

Figure B- 9: Design B-KNN Normalized Confusion Matrix (1 Day PP)

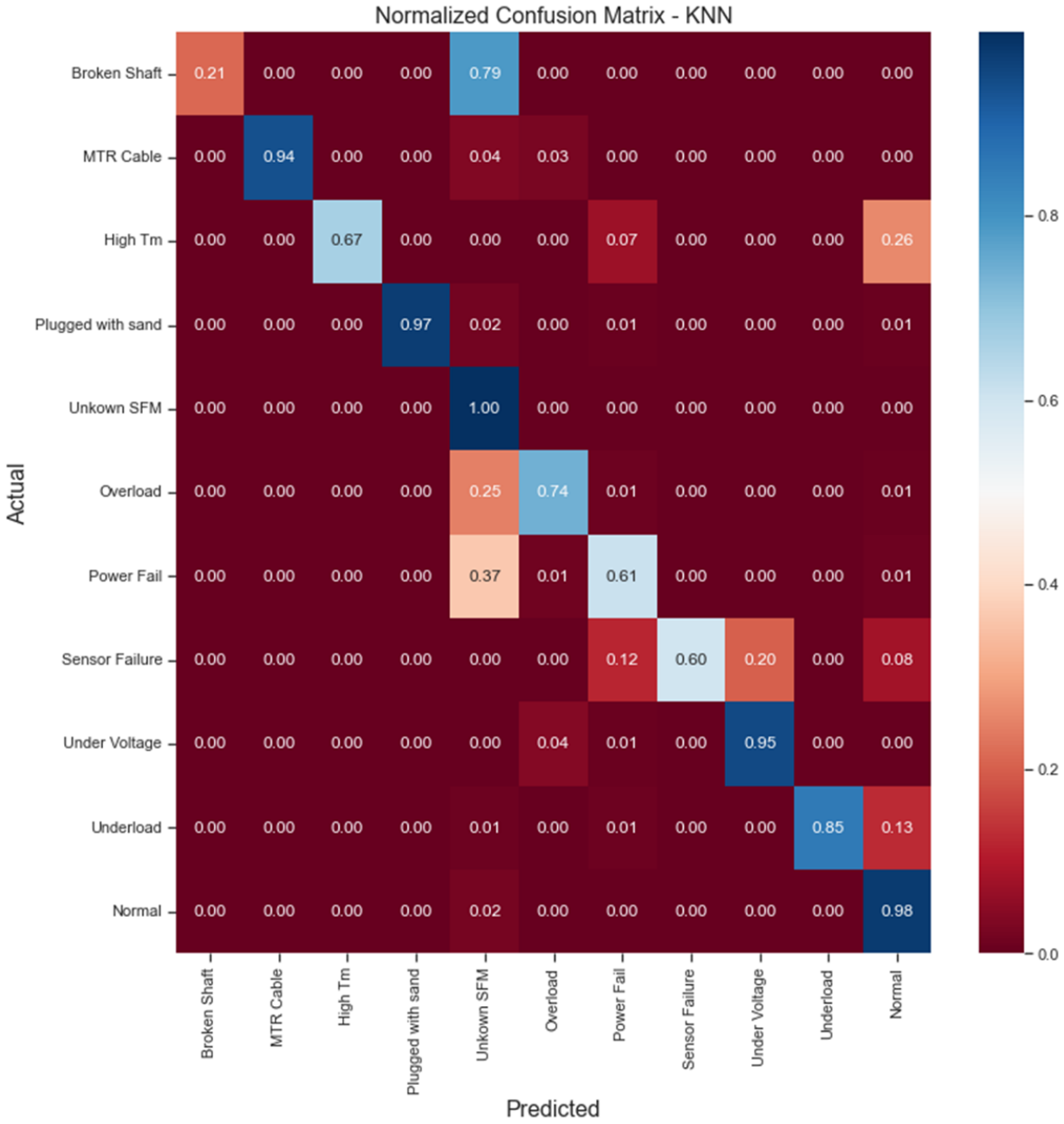

Figure B- 10: Design B-KNN Normalized Confusion Matrix (3 Hours PP)

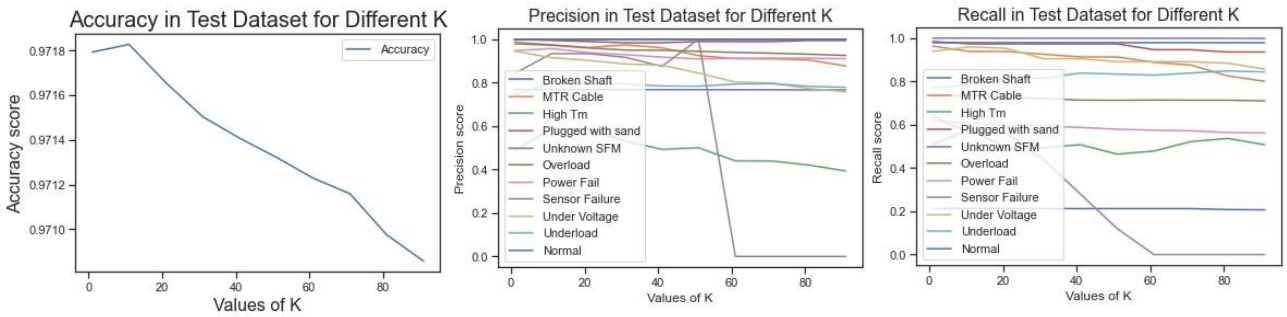

Figure B- 11: Design B-K Values Determination for KNN Model (3-hours PP)

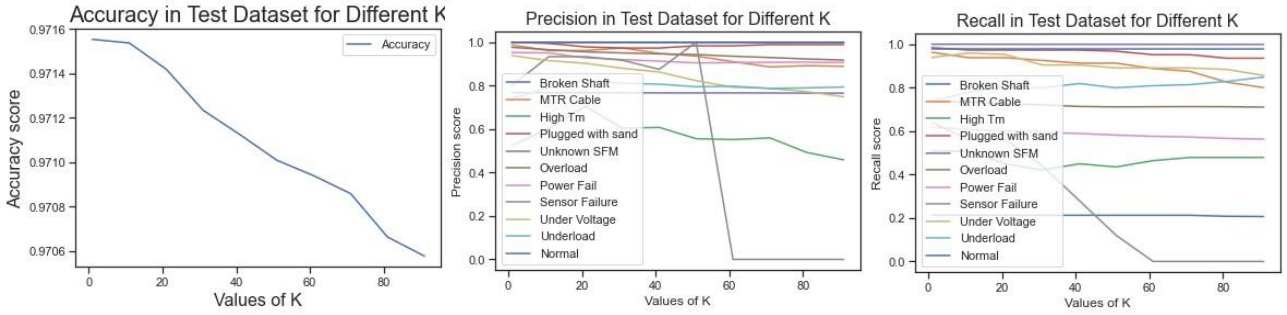

Figure B- 12: Design B-K Values Determination for KNN Model (1 Day PP)

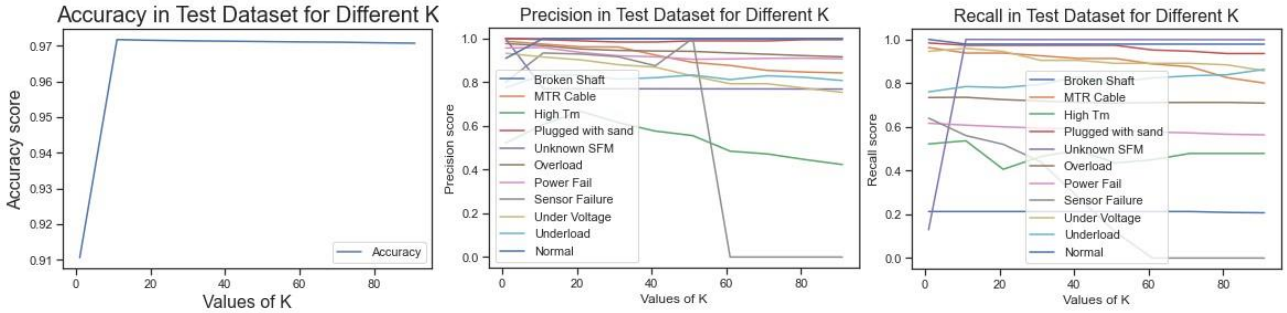

Figure B- 13: Design B-K Values Determination for KNN Model (3 Days PP)

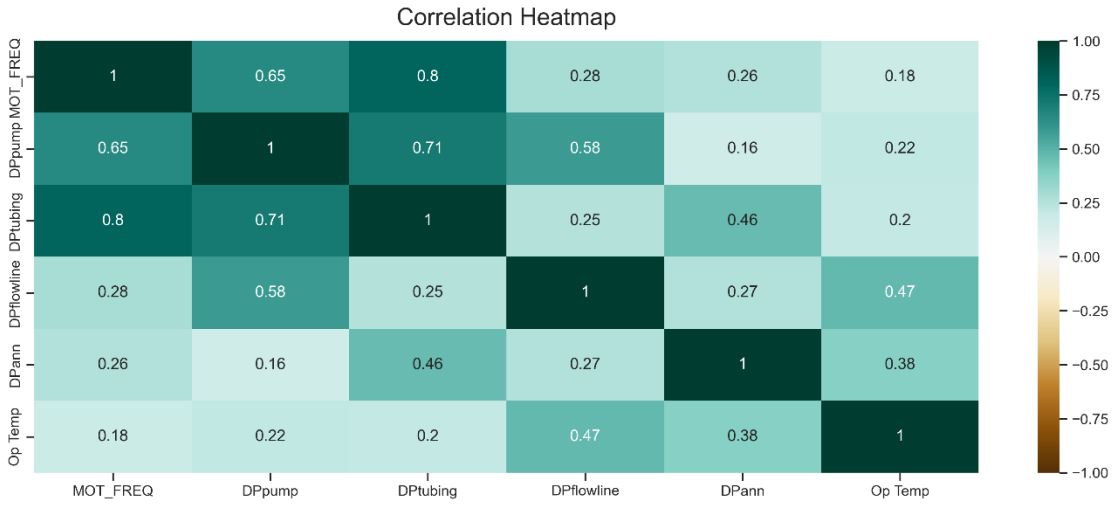

Figure B- 14: Design B-Calculated Variables Correlation Heatmap Matrix

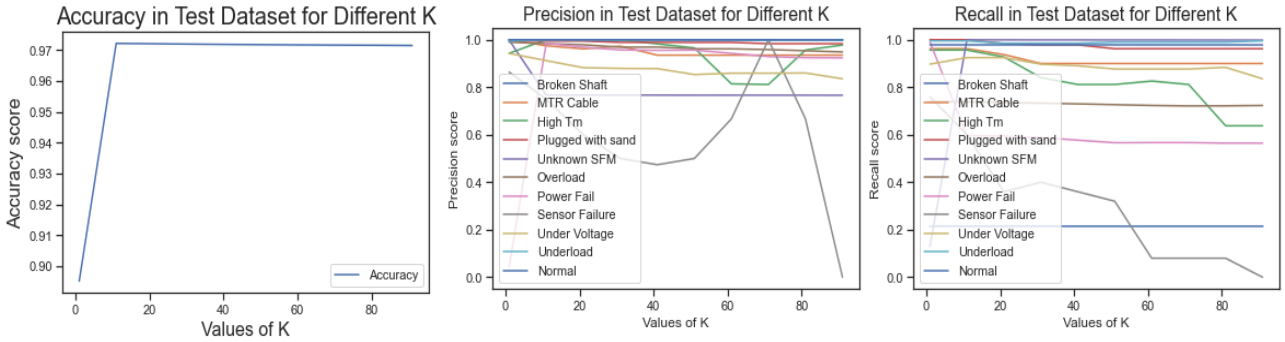

Figure B- 15: Design C-K Values Determination for KNN Model (3-hours PP)

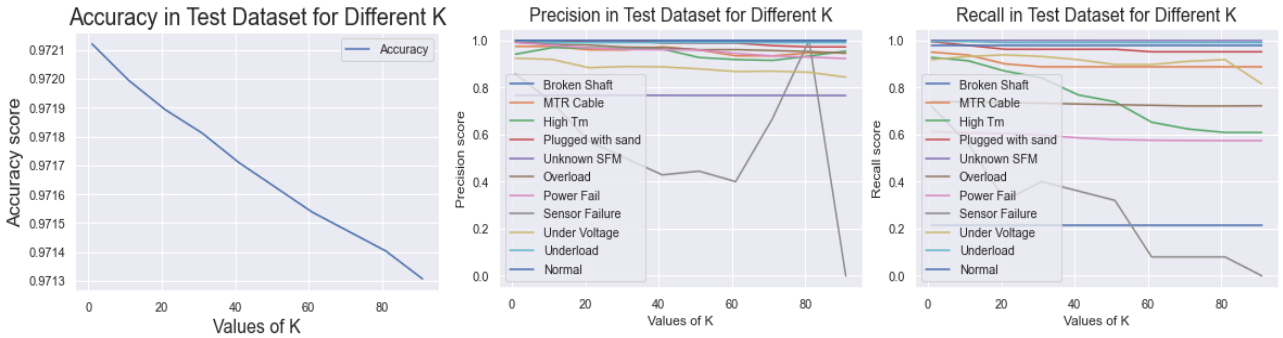

Figure B- 16: Design C-K Values Determination for KNN Model (1 Day PP)

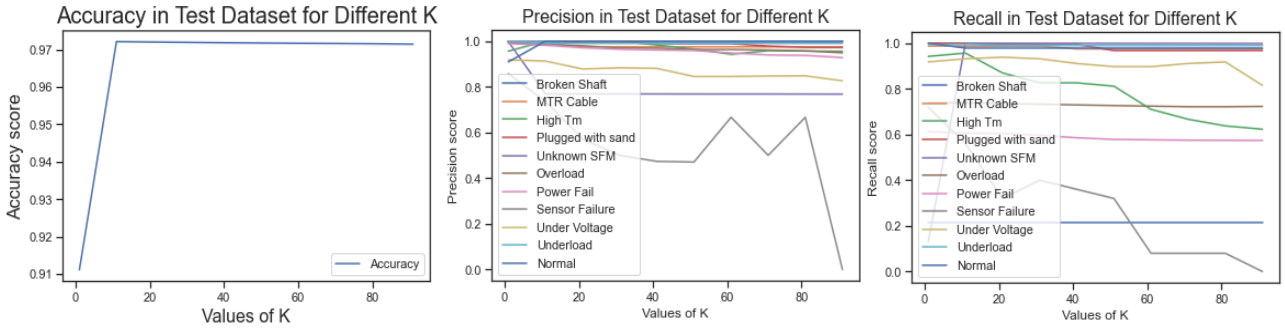

Figure B- 17: Design C-K Values Determination for KNN Model (3 Days PP)

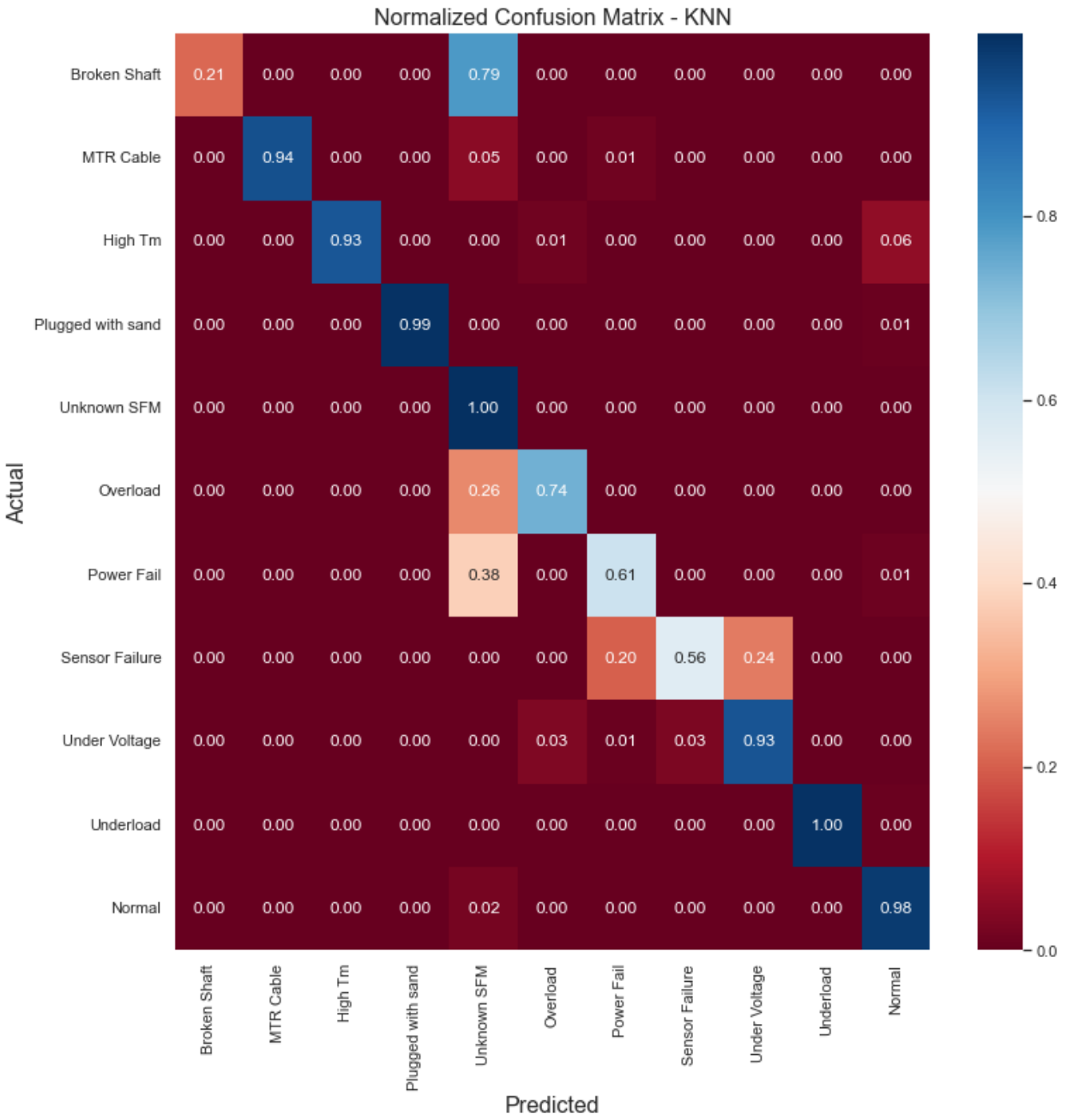

Figure B- 18: Design C-KNN Normalized Confusion Matrix (1 Day PP)

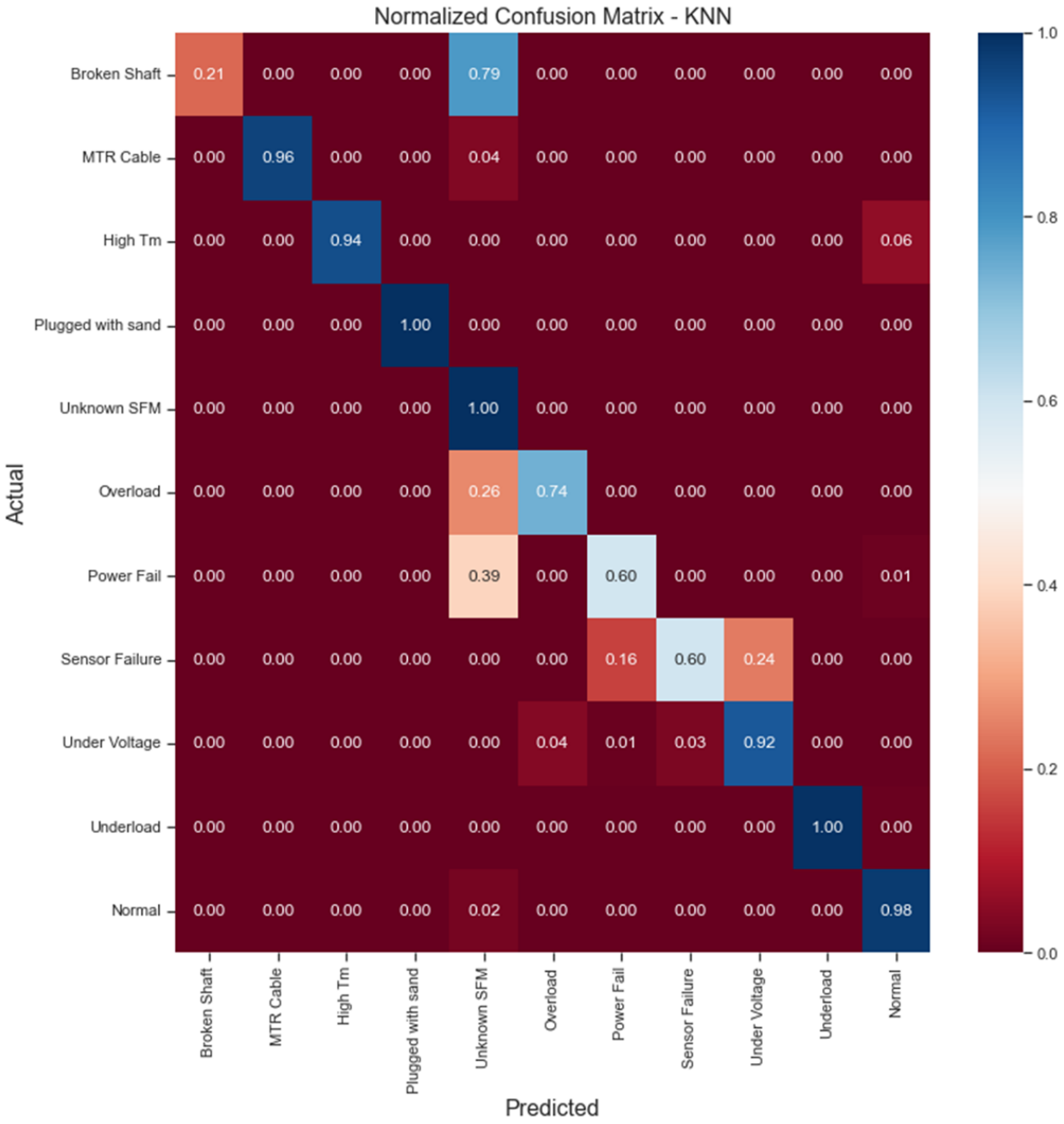

Figure B- 19: Design C-KNN Normalized Confusion Matrix (3 Hours PP)

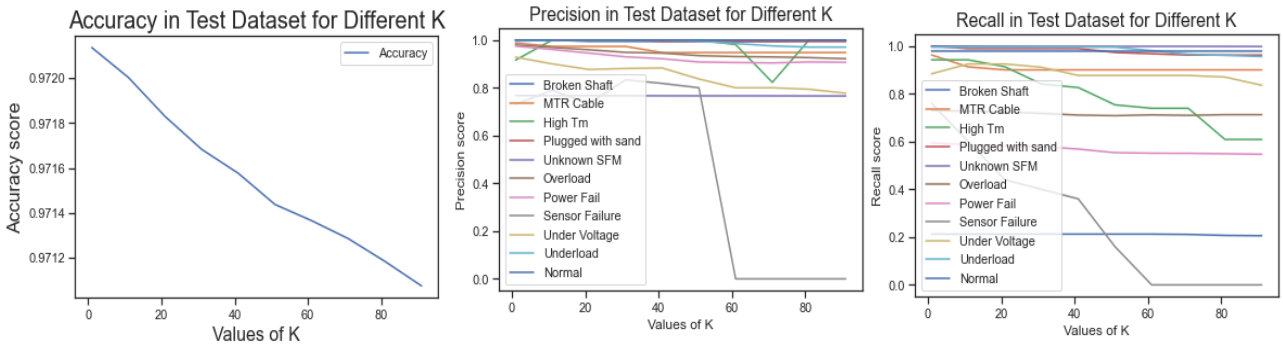

Figure B- 20: Design D-K Values Determination for KNN Model (3-hours PP)

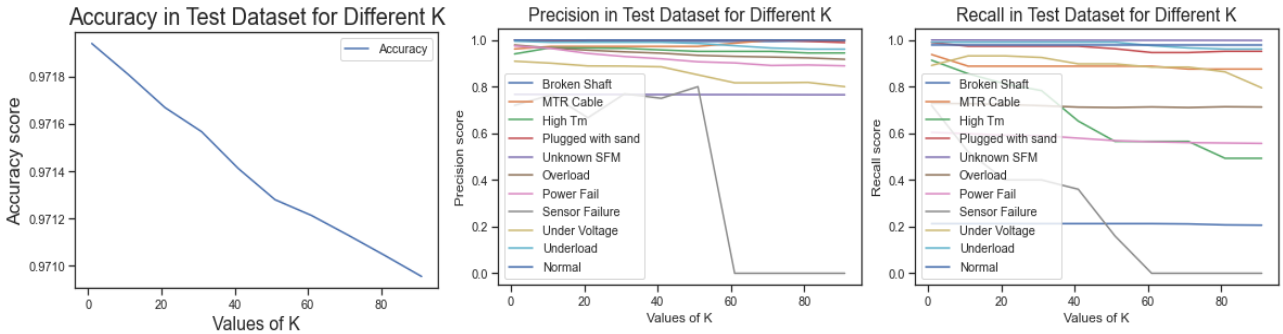

Figure B- 21: Design D-K Values Determination for KNN Model (1 Day PP)

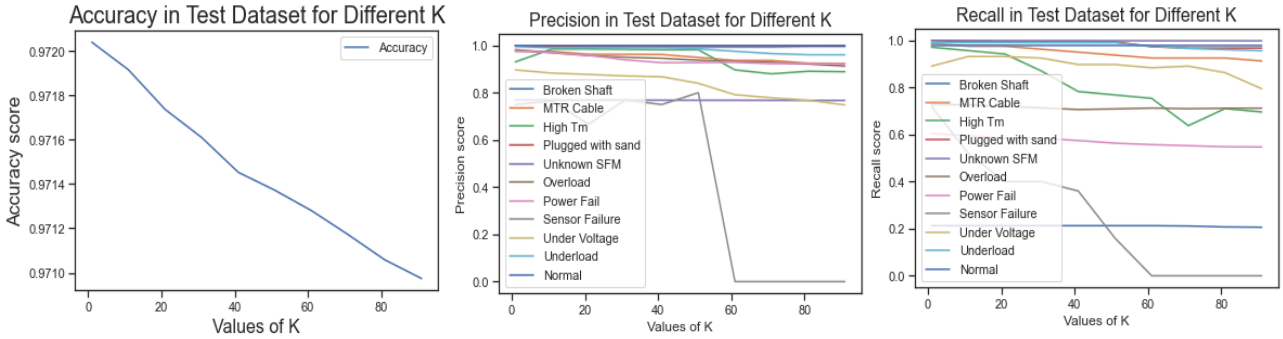

Figure B- 22: Design D-K Values Determination for KNN Model (3 Days PP)

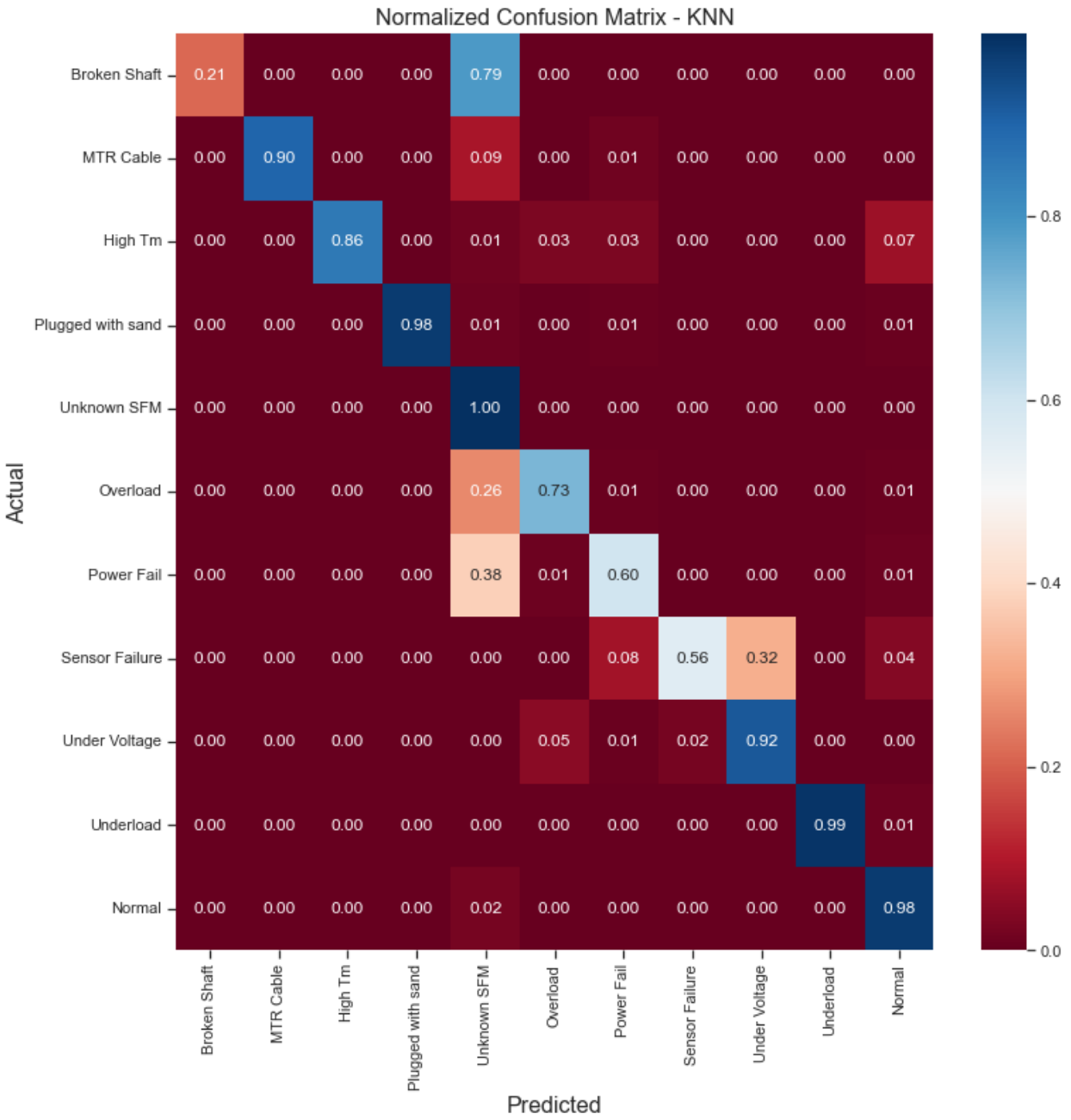

Figure B- 23: Design D-KNN Normalized Confusion Matrix (1 Day PP)

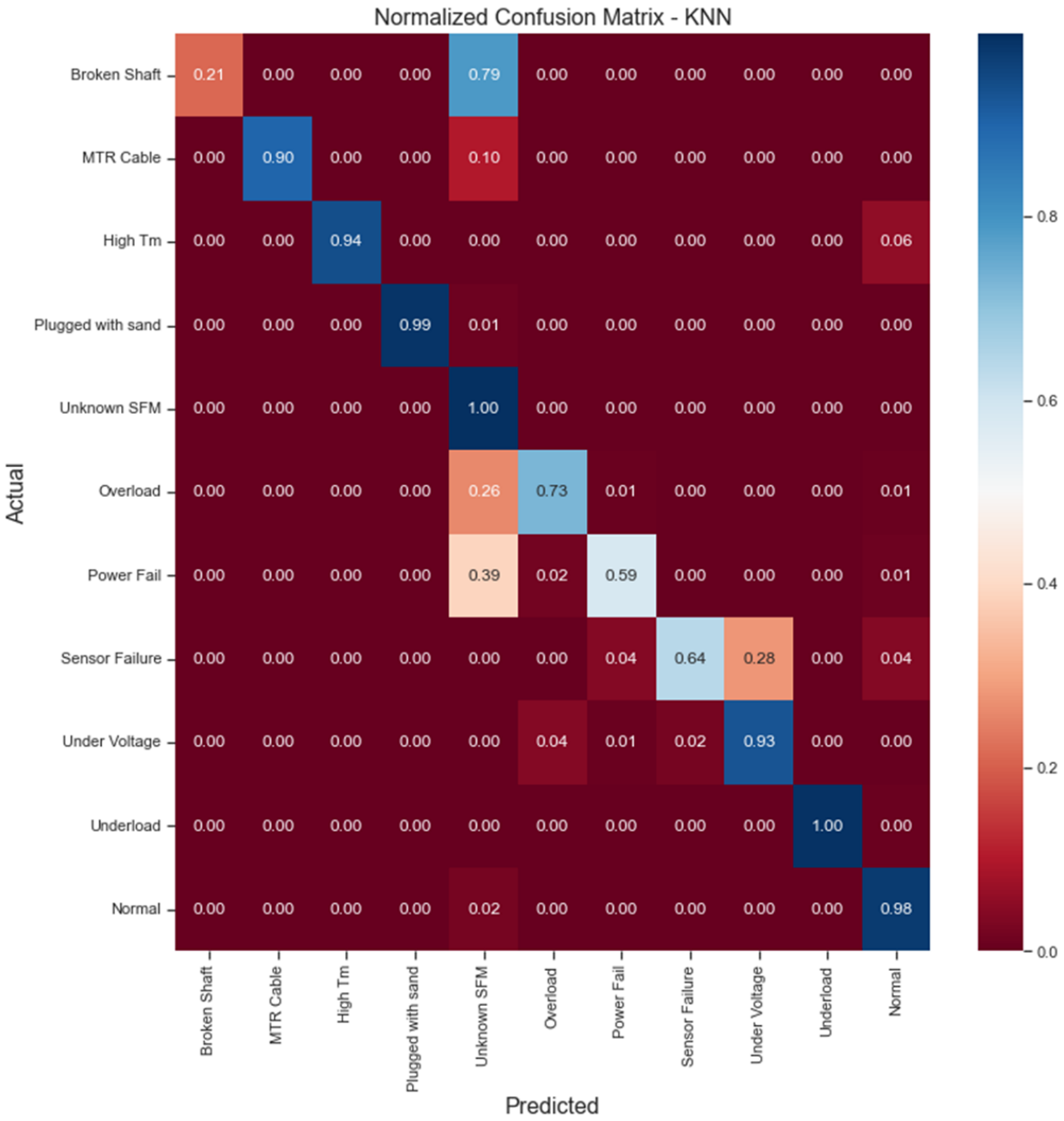

Figure B- 24: Design D-KNN Normalized Confusion Matrix (3 Hours PP)

Appendix C

Table C- 1: Design B- Hyperparameters Optimization Results

ML Models	Hyperparameter to be Tuned		Tuned Hyperparameter
	Hyperparameters	**Different Values**	
RF	max_samples	[1000, 2000, 3000, 4000, 5000]	5000
	n_estimators	[200, 500, 1000, 2000]	200
	max_depth	[3,4,5,6]	6
	min_samples_split	[5,10,15]	5
DT	max_leaf_nodes	[1, 2, 3, 4, 5, 6, 7, 8, 9, 10]	10
	min_samples_split	[9, 10]	9
MLP	hidden_layer_sizes	[(50,50), (50,100), (100,1)]	(50, 50)
	Activation Function	[relu, tanh, logistic]	Relu
	Alpha	[0.0001, 0.05]	0. 0001
	learning_rate	[constant, adaptive]	adaptive
KNN	K values	From 1 to 100	10
GNB	var_smoothing	From 1 to 10^{-9}	1.0
LR	C	[0.001,0.01, 0.1, 1.0, 10]	0.001

Table C- 2: Design E- Hyperparameters Optimization Results

ML Models	Hyperparameter to be Tuned		Tuned Hyperparameter
	Hyperparameters	**Different Values**	
RF	max_samples	[50,100, 200, 500]	500
	n_estimators	[100, 200]	100
	max_depth	[3,4,5,6]	6
	min_samples_split	[5,6, 7]	5
DT	max_depth	[1,3,5,7,9,11,12]	8
MLP	hidden_layer_sizes	[(50,50), (50,100), (100,1)]	(50, 50)
	Activation Function	[relu, tanh, logistic]	tanh
	Alpha	[0.0001, 0.05]	0.0001
	learning_rate	[constant, adaptive]	Constant
	Solver	Adam	Adam
KNN	K values	From 1 to 100	5
SVM	C	[0.1, 1, 2, 5]	5
	gamma	[0.001, 0.01, 0.1, 1, 10]	10
	kernel	[rbf, linear]	rbf
LASSO	Alpha	1e-08 to 1e+08	0.0001
Ridge			0.1
Polynomial	Degree	[2, 3, 4, 5, 6]	2

Table C- 3: Designs C & D- Hyperparameters Optimization Results

ML Models	Hyperparameter to be Tuned		Tuned Hyperparameter	
	Hyperparameters	Different Values	Design C	Design E
RF	max_samples	[1000, 2000, 3000, 4000, 5000]	5000	5000
	n_estimators	[200, 500, 1000, 2000]	500	2000
	max_depth	[3,4,5,6]	6	6
	min_samples_split	[5,10,15]	5	5
DT	max_leaf_nodes	[1, 2, 3, 4, 5, 6, 7, 8, 9, 10]	10	10
	min_samples_split	[9, 10]	9	9
MLP	hidden_layer_sizes	[(50,50), (50,100), (100,1)]	(50, 100)	(50, 50)
	Activation Function	[relu, tanh, logistic]	Relu	Relu
	Alpha	[0.0001, 0.05]	0.0001	0.0001
	learning_rate	[constant, adaptive]	Constant	Constant
	Solver	Adam	Adam	Adam
KNN	K values	From 1 to 100	10	10