IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received April 27, 2021, accepted May 12, 2021, date of publication May 21, 2021, date of current version June 2, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3082557

Deep Learning and Time-Series Analysis for the Early Detection of Lost Circulation Incidents During Drilling Operations

MOHAMMAD ALJUBRAN, JOTHIBASU RAMASAMY, MOHAMMED ALBASSAM, AND ARTURO MAGANA-MORA

Drilling Technology Team, EXPEC Advanced Research Center, Saudi Aramco, Dhahran 31311, Saudi Arabia

Corresponding author: Arturo Magana-Mora (arturo.maganamora@aramco.com)

ABSTRACT Drilling operations consist of breaking the rock to deepen a wellbore for oil or gas extraction. A drilling fluid, circulating from the surface through the drill pipe and from the annulus to the surface, is used to remove rock cuttings and maintain hydrostatic pressure. Drilling fluid lost circulation incidents (LCIs) are major sources of non-productive time (NPT) in drilling operations. These incidents occur due to preexisting natural fractures (vugs, caverns, etc.) and/or drilling-induced hydraulic fractures. The initiation of an LCI could lead to other hazardous drilling phenomena, such as formation influx or kick/blowout, stuck pipe incidents, among others. LCIs are typically monitored at the rig site by observing drilling fluid levels in the fluid tanks. This manual process incurs missing the occurrence or late detection of LCIs. Machine learning (ML) and deep learning (DL) classification algorithms are powerful in processing time-series data and achieving early detection of such temporal phenomena. In this study, we performed a large-scale analysis of the surface drilling and rheology data obtained from historical wells with LCIs. This analysis includes primary and secondary preprocessing steps including, aggressive sampling, feature engineering, and window normalization to derive generalizable DL models for real-time operations. Focal loss was utilized to account for data class imbalance and train robust and generalizable models. The results obtained from different ML/DL algorithms showed that one-dimensional convolutional neural network models resulted in the best performance with state-of-the-art precision, recall, and F_1 scores of 87.34%, 73.40%, and 79.77%, respectively, on unseen test drilling data.

INDEX TERMS Circulation losses, deep learning, drilling operations, industrial applications.

I. INTRODUCTION

A. BACKGROUND

Rapid industrialization across the globe has significantly increased the energy demand over the past few decades. Subsequently, the oil and gas industry has been drilling and producing hydrocarbon at growing trends to meet this demand. Upon the depletion of readily accessible reservoirs, exploration and exploitation of hydrocarbon reserves in complex geological formations have become more common. Consequently, the oil and gas drilling industry has been performing operations in more challenging environments, e.g., deep-water, horizontal extended reach, among others.

The associate editor coordinating the review of this manuscript and approving it for publication was Utku Kose.

Reserves are hydrocarbon-bearing rocks located deep underground, typically from a few to several thousand feet below the surface. Drilling is an operation that breaks rock from the surface to reach the target hydrocarbon reservoirs. Drilling involves the penetration through different types of formations with different lithology, porosity, permeability, etc. In the example shown in Fig. 1, five different formations have to be drilling before reaching reservoir formation.

Drilling a well consists of three major operations: drilling, casing, and cementing. These operations are conducted for the drilling of each section or formation. For instance, the entire sequence of drilling, casing, and cementing is completed for formation A before starting the drilling of formation B. This is due to the difference in the nature of the formation, mud density and property requirements, etc. Temperature and pressure naturally increase with the increasing

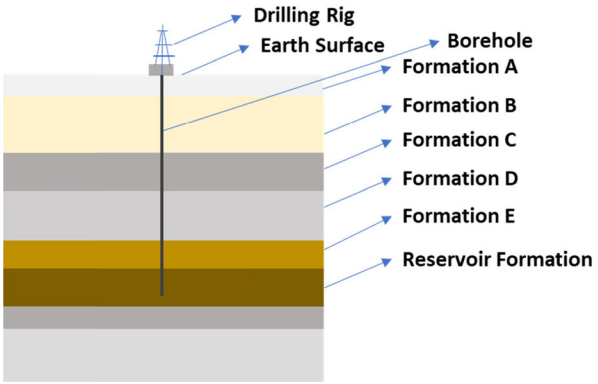

FIGURE 1. Oil and gas drilling of a vertical well across different rocks or formations.

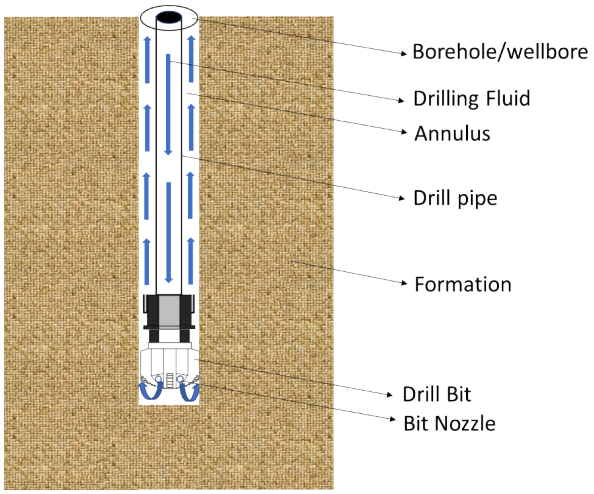

FIGURE 2. Demonstration of drilling fluid circulation path during drilling operations.

depth of the wellbore. During drilling, the pressure inside the borehole is less than that of the formation around it. The difference in pressure would lead to wellbore instability. In order to stabilize the wellbore, a fluid (referred to as drilling fluid or drilling mud) of a certain density is used. Drilling fluid is circulated from the surface through the drill pipe, exits through drill bit nozzles, and flows through the annulus (space between the drill pipe and wellbore) back to the surface, as shown in Fig. 2.

Drilling fluid is utilized to carry out several functions during drilling. The most important function is to maintain wellbore stability by providing hydrostatic pressure in the hole to balance formation pressures. Therefore, the borehole has to always be filled with fluid for the stability of the wellbore. Moreover, the density of the fluid is tuned carefully to keep the pressure at the bottom of the borehole slightly higher than that of the formation pressure. In addition to providing hydrostatic pressure to balance the formation pressure, drilling fluids serve several other functions, e.g., carrying drilled cuttings to the surface, cooling, and lubricating downhole equipment. Based on the type of liquid carrier used to formulate the drilling fluid, these fluids are classified into three major categories: water-based, oil-based, and synthetic-based fluids. There are several chemicals (additives) added to the drilling fluid system to carry out the functional tasks. Essential additives include viscosifiers, fluid loss control additives, and weighing materials. Depending on the nature of formations and drilling requirements, other additives are also mixed in the drilling fluid system. For instance, when drilling through water-sensitive shale/clay formation, clay inhibitors/stabilizers are commonly added. Meanwhile, highly permeable zones are drilled with drilling fluids having bridging materials with relatively larger grain sizes to minimize the loss of fluid into the formation.

Drilling operations represent a is a dynamic and challenging environment due to natural, mechanical, and human factors, where several issues could be encountered, which may lead to non-productive time (NPT). Drilling fluid lost circulation incidents (LCIs) are among the major contributors to drilling NPT [1] that can increase the operational costs substantially, especially in offshore operations [2]. The loss of drilling fluid could be partial or total, depending on the nature of the formation being drilled. In cases where circulation losses are minimal, the drilling operation may continue with partial returns of mud to the surface and under necessary precautions to ensure safe drilling operations. However, when circulation losses are severe, the drilling operation has to be stopped, and losses must be cured to regain circulation [3], [4].

In LCIs, the fluid in the borehole may be lost into the formation. Consequently, bottom-hole pressure may be insufficient to balance formation pressure and prevent formation from collapsing into the wellbore. When drilling reservoir sections (depths where hydrocarbon is trapped), lost circulation may lead to an uncontrolled influx of hydrocarbon into the wellbore, referred to as kick or blowout depending on their severity [5]. In addition, wellbore collapse may occur due to reduced pressure in the annulus resulting from lost circulation, leading to drill pipe and downhole tools being buried in the hole [6]. Loss of annular pressure may result in stuck pipe incidents [7], [8], which would substantially increase the complexity of mitigation operations and hence NPT. Globally, the cost of lost circulation NPT is estimated to be US$ 2–4 billion yearly [1].

Loss of circulation occurs while drilling through highly permeable, fractured, and vugular formations, seen in Fig. 3. The magnitude of the drilling mud losses depends on the type of loss zones present in the formation. Seepage loss of circulation is experienced while drilling highly permeable zones. With adequate lost circulation material (LCM) added to the drilling fluid system, drilling through these zones could be continued. Generally, the extent of losses and complexity of mitigation increase with wider, deeper, and interlinked fractures. Natural fractures can be present in any formation, but they occur mostly in geological settings with ongoing tectonic activity. Vugs and cavities are present mostly in carbonate formations [9]–[11].

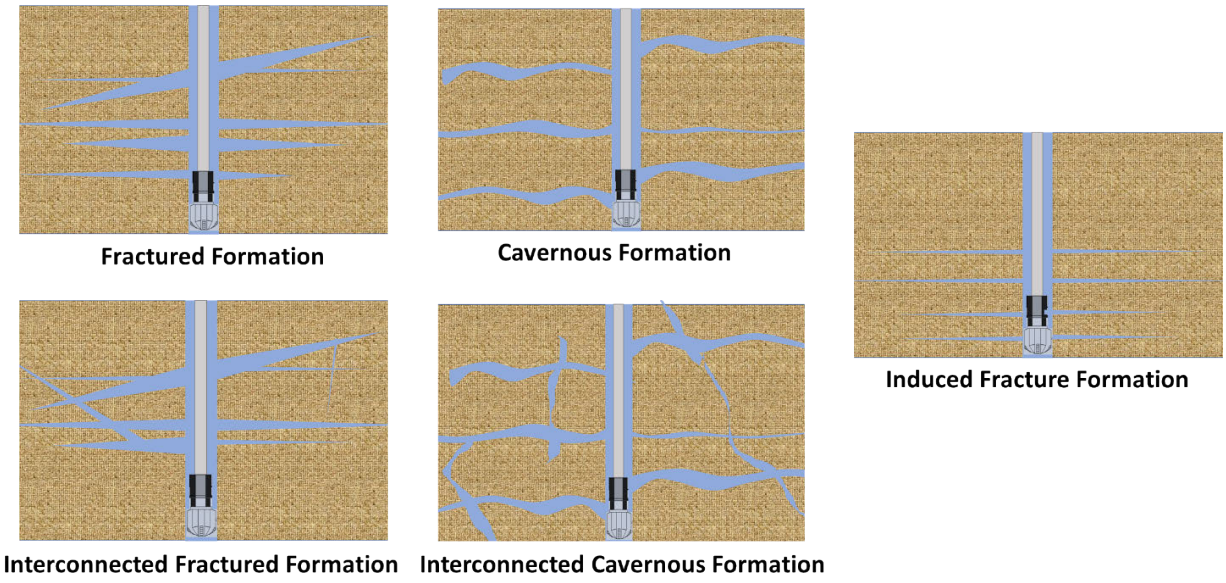

FIGURE 3. Types of Earth formations that could cause drilling mud circulation losses when encountered.

Induced lost circulation occurs due to inducing new fractures while drilling, which often happens while drilling weak formations with higher mud weights. The high overbalance of the mud is enough to induce new fractures, or widen the already existing fractures, in weak formations while drilling [12]. Induced mud losses occur by hydraulic fractures extending from the wellbore deep into the formation. Induced lost circulation is often encountered when drilling formations having narrow mud weight windows, which is the safe mud weight margin between formation pressure and fracture pressure. These induced losses could be mitigated relatively quickly as the fractures are usually not deep nor interconnected.

Various lost circulation scenarios associated with narrow mud weight margins are weak and/or depleted zones, deep-water formations, formations with natural fractures, and deviated wellbores [13], [14]. Reducing the pore pressure in weak and depleted formations can cause a significant reduction in the fracture pressure, therefore, lowering the pressure-bearing capacity. In deep-water environments, considerable depth of water can result in lowering the formation fracture pressure, leading to a narrow mud weight window. In such situations, it is very challenging to maintain the required wellbore pressure. Sudden and small fluctuations in wellbore pressure may result in swab and surge, which can further complicate operations. In the case of deviated wellbores, the higher the inclination angle of the wellbore, the narrower the mud weight window. This could even lead to a ''zero'' mud weight window and an un-drillable section. Naturally, existing fractures in the formation can significantly affect the pressure-bearing capacity of the wellbore. A tiny fracture in the formation may result in lower tensile strengths near the wellbore and, subsequently, the fracture will propagate when the wellbore pressure overcomes the near-wellbore hoop stress. When the fracture is wide and deep, the maximum pressure a wellbore can withstand may be reduced to the value of minimum principal in-situ stress or pore pressure [15].

LCIs are categorized broadly into three major tiers based on the number of barrels (bbls) of drilling mud lost.

Seepage LCI: < 40 bbls/hr

Partial LCI: > 40 bbls/hr and < 100 bbls/hr

Severe/Total LCI: > 100 bbls/hr

Various types of LCMs (e.g., particulates, fibrous materials, flaky materials, sized bridging materials, among others) are used for mitigating lost circulation [16]–[18]. For seepage loss zones, the drilling fluid is loaded with the above-mentioned materials in particular proportions based on the porosity and permeability of the formation, and the drilling operation is not ceased. In partial and severe/total loss scenarios, the drilling operation has to be stopped and the LCM slurry is prepared with the above-mentioned materials and pumped downhole to mitigate loss by plugging and bridging. The LCM slurry might fail to mitigate losses in the case of severe/total losses. In such cases, different types of polymers, resins, cement plugs, etc., could be used to cure losses and regain mud circulation.

B. LITERATURE REVIEW

Studies on early identification of LCIs and mitigation methods have been at the center of attention in drilling for oil and gas wells for some time as these events may result in wellbore instability and additional costs due to the lost drilling fluid. Equivalent circulating density (ECD) is a physics-based model that has been consistently used to avoid LCIs. As a result, several studies have been devoted to developing annular pressure loss models to predict ECD and equivalent static density (ESD). The primary objective of utilizing an

ECD model is to ensure that the applied static/dynamic fluid pressure is within the drilling margin. That is, to ensure a drilling mud having sufficient mud weight to avoid wellbore instability but not greater than the fracture pressure that would induce formation fractures. Induced formation fractures and, therefore, mud losses occur when the bottom hole pressure (BHP) rises above fracture pressure.

Ahmed and Miska [19] developed equations to calculate annular frictional pressure for both power-law and yield power-law fluids traveling across the annulus. Their equations consider various flow regimes (i.e., laminar, transitional, and turbulent), inner and outer diameter ratio, and eccentricity. However, there is a discrepancy in the proposed equation as it assumes a clean drilling fluid passing across the annulus, which may be invalid considering the fact that one of the primary functions of the drilling fluids is to carry and remove cuttings from the bottom-hole to the surface. Therefore, these drilled cuttings impact the rheology, density, and velocity of the drilling mud, thus, increasing annular pressure loss. Whenever ECD goes above fracture pressure, then the rock formation can break and initiate mud loss. In order to address the assumptions of having a clean drilling fluid, Bassam [20] further developed Ahmed and Miska's equation to include the transport of drill cuttings in the annulus by coupling a pressure gradient equation with a cuttings transport model to predict ECD. The model can be used to flag an increase in ECD and alert the driller if the ECD goes beyond the fracture pressure at any point of time and at different depths, not only at total depth. This allows the crew to adjust drilling surface parameters or mud parameters to ensure that the applied static/dynamic fluid pressure is within the mud weight margin and to avoid induced loss circulation incidents.

Although physics-based models can be used to avoid LCIs, data-driven models offer another layer of information that can be used to estimate or predict unseen events based on historical data from offset drilling operations as well as real-time data collected during drilling operations. In this direction, different machine learning (ML) and deep learning (DL) models have been developed to predict mud loss of circulation from surface parameters [21]–[24]. Moazeeni *et al.* [25] reported one of the earliest studies utilizing ML to develop a model capable of predicting LCIs in different areas of a specific oilfield, as well as to estimate the quantity and quality of LCIs. The authors restricted their study to only a single formation that was often associated with severe losses. The data was collected from 32 drilled wells, using their daily drilling reports, and filtered based on location (Northing and Easting). The study incorporated several inputs for the model, including depth of the incident, length of open hole section, drilling bit size, mud weight, mud rheological parameters, volume of drilling mud losses, among others. More recently, another study [26] developed a model that predicts LCIs using support vector machines with radial basis function derived from six drilling parameters (mud flow pump, rate of penetration, pipe rotation, standpipe pressure, torque, and weight on bit) and the depth of the lost circulation zone obtained from two wells with LCIs. The data from the first well was utilized for training, whereas the second well data was used for validation.

While several ML/DL models for the detection of lost circulation were reported, the lack of generalization capabilities and required accuracy of the models have restricted their implementation to guide real-time operations [27]. Arguably, the scarce data from the limited number of considered wells with LCIs may not be enough to represent all conditions and characteristics leading to an LCI. Training competent ML/DL models requires a dataset that captures the joint distribution of the input and output variables. This is challenging in drilling as various parts of the joint distribution only occur in specific aspects of the drilling operations. In addition, drilling data also capture other dimensions of variability associated with different types and setups of rig structures, sensory equipment, down-hole formation properties, among others. Additionally, drilling data are typically generated from multiple sources. These sources or sensors could also belong to different drilling rig contractors and service providers who collaborate to complete a variety of tasks and operations. As a result, drilling databases host a collection of records stemming from multiple sources, which results in asynchronous data (the clock in every microcontroller is slightly different) and potentially inconsistent sensor sampling frequency [28, 29]. Whereas these problems could be resolved or alleviated by transmitting a common pulse to sensors to collect records simultaneously, this is not a standard practice in drilling operations.

Moreover, drilling is an aggressive process with significant levels of vibration. It also involves a variety of sensors stemming from different sources, raising the potential of electromagnetic interference. This is further exacerbated by the rough drilling environments (desert, forests, sea, etc.) where power fluctuation is common. On-demand sensor calibration is also a must (e.g., hook load is measured using a load cell, which needs to be recalibrated on-demand for each newly run bottom-hole assembly throughout the drilling operation). Due to the absence of on-demand and continuous sensor maintenance and recalibration, these challenges result in inaccurate drilling data associated with random noise, local outliers, and mis-calibrated records.

In this study, we identified a statistically representative number of wells from different fields and extracted the surface drilling parameters that capture the different conditions leading to an LCI. This study focuses on data preprocessing, feature engineering, and normalization, as well as model evaluation to derive a robust and generalizable DL model. Additionally, the methodology proposed in this study formulates the classification task as a time-series problem to accurately detect the trends of the parameters instead of the highly variable readings from the surface sensors. The next section describes the methodology followed for data extraction, normalization, and training of the ML/DL models.

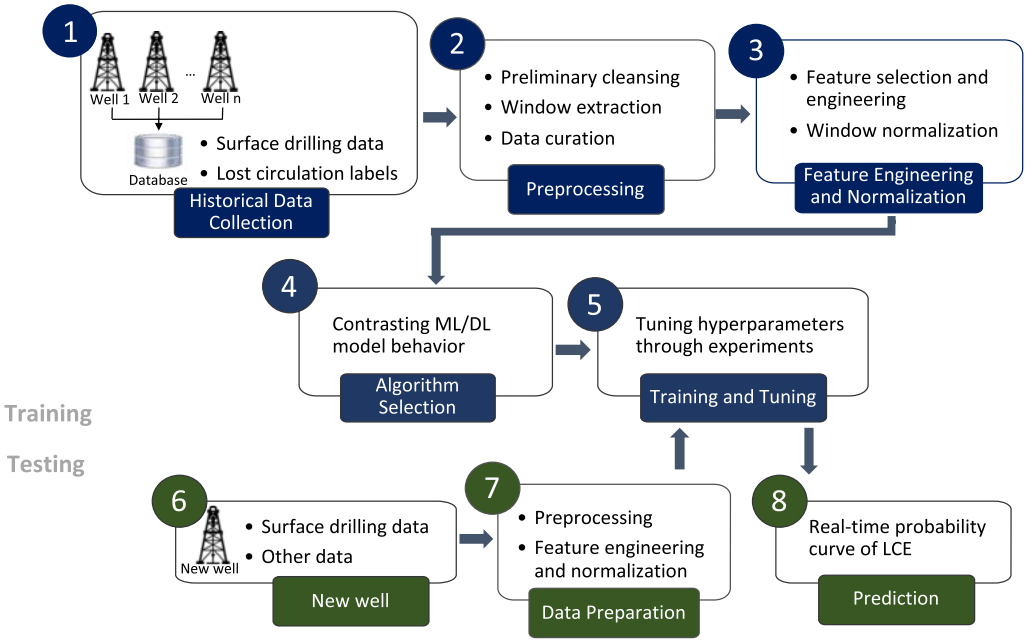

FIGURE 4. General demonstration of the proposed ML/DL model development approach can be divided into two tiers: training (blue) and testing (green). Training starts with historical drilling data collection, preprocessing, feature engineering and normalization, algorithm selection, and training and tuning. Testing involves introducing a blind historical well or real-time well (one that has not been shown to the model during training and validation) to the model where data are first prepared in a similar fashion to the preceding training phase, then run through the model to generate real-time probability curve indicating the likelihood of LCI at a given point of time.

II. METHODS

The proposed DL model development approach for early LCI detection goes through two major phases: training and testing, as shown in Fig. 4. The training phase requires the collection of sensors data obtained from historical drilling operations with LCIs, preprocessing and time series formulation, curation, feature engineering and selection, data normalization, model selection, and model hyperparameter tuning. The testing phase, on the other hand, refers to the deployment of the already trained and validated model to be tested on unseen wells (either historical or real-time). In the testing phase, data are handled and prepared in a similar fashion to that of the preceding training phase before they are fed into the DL model. The deployed model generates a probability curve indicating the likelihood (i.e., between 0 and 1.0 referring to lowest and highest probability, respectively) of encountering an LCI in the corresponding drilling operations.

The following subsections introduce the techniques and algorithms utilized for data collection and preprocessing, feature engineering, data normalization, and model training and evaluation.

A. DATASET

This study is based on historical data collected during drilling operations using standard drilling rig equipment and apparatus [30]. These data contain the stored values of the multiple sensors used by the rig crew to guide operations in real-time. Although this data is stored, not all data are relevant to derive ML models. Consequently, data collection and processing are a crucial steps for deriving generalizable ML/DL models. To alleviate the challenges associated with drilling data, it is essential to collect data extracted from a significantly large number of LCIs from multiple wells. These wells have to statistically represent the signals leading to LCI while being able to discriminate from noise, outliers, missing, and irrelevant data as well as to identify the different earth formations at different depths that may lead to LCIs.

In this study, we identified ∼200 wells with severe or total losses from multiple fields and extracted the surface sensors data (i.e., drilling parameters) with respect of time with an average frequency of 0.2 hertz per record (i.e., one data vector every five seconds). Table 1 shows the list of the selected surface sensors data with their respective units. Clearly, the differences between flow-in rate (FLWIN), i.e., the fluid entering the wellbore through the drillpipe, and flow-out rate (FLWOUT), i.e., the amount of fluid exiting the wellbore, could be used to directly identify early signs of LCIs. However, the readings of these two sensors are in different units, have time lags, and contain large uncertainties due to the sensor technology. Consequently, we also considered weight on bit (WOB), and revolutions per minute (RPM) as these are the variables controlled by the driller and which have a significant effect on the rest of the parameters (i.e., torque – TQ, rate of penetration – ROP, standpipe pressure – SPP, and hook load – HKL). The trends and linear or nonlinear relationships of these data can be used to identify the signals leading to LCI.

TABLE 1. Selected surface drilling parameters.

Parameter	Units
Weight on bit (WOB)	Kilo pound (klbf)
Hook height (HKHT)	Feet (ft)
Hook load (HKL)	Kilo pound (klbf)
Torque (TQ)	Pound-foot (kft.lbf)
Stand pipe pressure (SPP)	Pounds per square inch (psi)
Flow-in rate (FLWIN)	Gallons per minute (gpm)
Flow-out rate (FLWOUT)	0-100%
Rate of penetration (ROP)	Feet per hour (ft/hr)
Revolutions per minute (RPM)	Rpm
Total mud system volume (PVT)	Barrels
Trip tank volume (TTV)	Barrels

For instance, a larger fracture may be surrounded by smaller interconnected cavernous formations, which would result in a sudden increase in ROP and a decrease in TQ prior to the LCI.

For each of the selected wells with LCIs, we extracted the surface drilling parameters described in Table 1. Since each well has roughly five million data vectors (surface parameters for collected during/months of drilling operations), only the relevant data must be selected. A window of 20 minutes surrounding the LCI (15 minutes before and five minutes after). This window is a 2-dimensional (2D) matrix where rows represent the 240 timesteps (in 20 minutes with a 0.2 Hz frequency) and the columns the drilling parameters, shown in Table 1. For each of the 20-minute segment of data surrounding an LCI, we used an additional 15-minute sliding window with a stride of one row (∼5 seconds) to produce the data samples assigned to the LCI category, i.e., each LCI is defined by 60 2D matrices of 15 minutes (180 rows) and 11 columns (drilling parameters and sensors readings).

In order to capture the normal drilling operations (i.e., not overlapping any LCI), we first extracted two 24-hour segments of data before the LCI while discarding 30 minutes of data between the two segments to avoid incorrectly capturing LCI patterns during this semi-automated labeling process. Additionally, we extracted 24 hours of data after the LCI. We then concatenated the three 24-hour segments of data (resulting in a 72-hour segment) and extracted several random 20-minute segments. We used the same sliding window approach for each of the different segments with capturing normal operations

Whereas we have one LCI per well, there exist many non-LCI or "normal" operations across these wells. If we simply sample one normal operation segment per LCI segment, then we only capture limited portions of the normal drilling operation probability distribution, which is extremely complex and multidimensional. Although class balance (equal number of LCI and non-LCI windows) is always desired in ML/DL model development, sampling one normal window per LCI window does not truly capture the complex

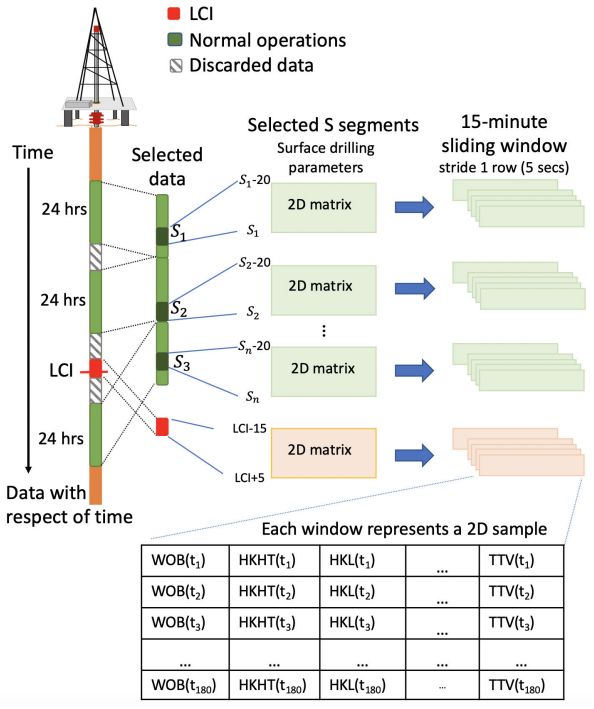

FIGURE 5. Methodology for the extraction of LCI and normal operation samples. Each well with LCI generates 180 non-LCI 2D samples and 60 LCI 2D samples. Each of these 2D samples is defined by a 180 × 11 matrix where columns represent the surface drilling parameters and rows the time steps in 15 minutes. The data surrounding the LCI are discarded to avoid incorrect labelling LCI patterns as normal operations.

drilling probability distribution function. Therefore, we performed aggressive sampling where we sampled three non-LCI windows for each LCI window. However, this results in class imbalance that may hinder ML/DL model training and requires a solution. This problem will be addressed using focal loss for training the ML/DL model, which is explained in detail in the following subsections. Fig. 5 shows the process for extracting LCIs and normal operations for a well. Note that we refer to the input of the model (2D matrix) as a sample. Each sample is defined by a 180 × 11 matrix, where columns represent the surface drilling parameters and rows the time steps in 15 minutes. Each well yields 240 samples based on a 1:3 aggressive sampling ratio of LCI:normal operations. In total, this process yields 48,000 samples (2D matrices) and 8,640,000 rows (1D vectors) for the 200 considered wells.

B. DATA PREPROCESSING

Data quality problems represent a significant challenge for developing predictive DL models, often the case when using drilling sensors data obtained at drilling rigs. Proper handling and preprocessing of drilling data are essential for DL algorithms to learn consistent and physically correct information from the data feed. The drilling literature addressed various aspects related to data quality, where preprocessing is essential [31]. Table 2 provides a summary of the general data preprocessing criteria conducted in the context of LCI.

TABLE 2. Data processing criteria.

Criteria	Description
Relevance	Drilling operations generate various forms of discrete and continuous data streams, hence, it is crucial to exclude sporadic and irrelevant features, which could hamper DL model training in latter stages.
Range check	Elimination or re-sampling data within soft upper and lower boundaries to eliminate extreme and physically meaningless drilling feature values (e.g., negative ROP).
Gap filling	Drilling data sometimes have one or more missing features across specific intervals of time, which must be either imputed or ignored during training. Whereas imputation is outside the scope of this work, wells with missing data across the intervals of interest are ignored during training. During evaluation, data processed through the model must be either complete or imputed a priori before utilizing the proposed model.
Duplication check	Drilling databases sometimes contain repeated or redundant features which hold perfectly or nearly identical values, hence these redundant features are checked and dropped at this stage. Note that redundancy is often encountered when the same physical quantity is measured using two different tools. For instance, in managed pressure drilling, FLWOUT is a Coriolis flow meter in addition to the standard rig flapper sensor located at the mud return line (a pipe through which mud flows from the well head to the mud tanks).
Unit standardization	Features stored in different measuring units (e.g., pounds versus kilo pounds) across different wells are standardized to be consistent and comparable in magnitude.

The next subsections introduce and discuss three specialized preprocessing techniques: 1) feature engineering, 2) window normalization, and 3) aggressive sampling. Unlike the preliminary preprocessing steps described in Table 2, these methods are more algorithmically advanced and capable of resolving complex data quality challenges, i.e., multiple data sources, inaccurate data, and data diversity.

1) FEATURE ENGINEERING

Feature engineering is a technique commonly used in ML/DL model development [32], [33]. This technique utilizes domain knowledge and expertise to extract a new representation of the raw data, aiming to improve the ML/DL model training process that involves nonconvex optimization over thousands or millions of parameters. There is an infinite number of features that could potentially be engineered for any application. The process of creating and selecting useful features heavily relies on domain knowledge, data structure (time-series in this application), and data quality issues [34]. In this study, we utilized four categories of features: 1) moving average filtering (MAF), 2) Gaussian filtering (GF), 3) differencing, and 4) drilling pumps on/off indicator.

MAF is a common representation of raw time-series records, which serves in denoising data and providing a zoomed-out snippet of a particular drilling activity. MAF has a single hyperparameter, n, which refers to the number of previous time-series data points. As in Eq. 1, MAF of a particular drilling feature can be computed at time t using the n preceding data points $\left[P_t, P_{t-1}, \ldots, P_{t-(n-1)}\right]$. Note that this simple MAF formulation weighs all data points equally [35].

$$MAF = \frac{1}{n}\sum_{i=t-(n-1)}^{t} P_i \tag{1}$$

We applied MAF twice for each feature with $n \in [24, 48]$ records, which correspond to $[2, 4]$ minutes, respectively, in the most common data frequency of 0.2 Hz. This allows for two levels of macroscopic trends and also two denoising stages, which the ML/DL model can learn from during training. MAF is useful in alleviating the problems of multiple data sources and inaccurate data. However, it is not ideal in denoising data as MAF is associated with a temporal lag, which may hinder early detection in classification tasks.

Whereas MAF is a causal filtering approach (only utilizes present and past data points), GF is a non-causal filtering approach, which does not result in temporal lag. Actually, causality is almost not relevant in this context since it does not involve forecasting the future. In this study, we utilized a one-dimensional GF to smoothen drilling features individually. GF is also a true low-pass, so it does not create high-frequency artifacts. For each feature x, one-dimensional GF convolves a one-dimensional kernel defined using a Gaussian distribution $g(x)$, as described in Eq. 2. Note that we assume a zero-mean Gaussian distribution in this application [36], [37].

$$g\left(x\right) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \tag{2}$$

GF is a linear filter that performs a weighted average giving more weight to the central timesteps and less weights to farther timesteps. These weights are defined by the standard deviation σ of GF. Larger σ values flatten the Gaussian bell curve, which in turn assigns more weight on neighboring timesteps and blurs the signal further. However, the noise type and frequency associated with drilling operations differ from one operation to another and from one sensory equipment to another. Whereas this type of noise is often random and time-variant, we tackled the noise frequency issue by engineering multiple GF features with different standard deviations, namely $\sigma \in [3, 24, 48]$. Overall, GF is useful in alleviating data inaccuracy, particularly noise, and local outliers.

Another engineered feature is based on a temporal derivative of each drilling parameter based on the first-order backward finite difference, as shown in Eq. 3, which estimates the derivative of a drilling parameter over time for point values P. We also parameterized this feature with d, the distance or timesteps between differentiated records. In this application, we used multiple $d \in [1, 24]$. This family of features is important to time-series analysis as it provides a sense of change with respect to time, which alleviates the multiple data sources issue of variant frequencies across operations

and sensory equipment.

$$gradient = \frac{P_t - P_{t-d}}{t - d} \tag{3}$$

We also curated an additional feature, based on technical domain knowledge of LCI, to indicate whether drilling fluid pumps are on or off. Early LCI signals are commonly associated with a sudden or gradual drop in FLWOUT due to loss of drilling fluid into the formation, followed by a sudden drop in FLWIN when the rig crew decides to turn off the drilling pumps. This behavior is similar to intervals where the rig crew simply turns on the rig pumps where both FLWOUT and FLWIN suddenly drop to zero (or other relatively low values in cases where sensors are mis-calibrated). Hence, we engineered an indicator feature that indicates local peaks and troughs in FLWOUT and/or FLWIN. This indicator takes one of three values: ''0'' if neither FLWOUT nor FLWIN have peaks nor troughs, ''1'' if either FLWOUT or FLWIN has a peak and/or trough, and ''2'' if both FLWOUT and FLWIN have peaks and/or troughs. This is done by firstly running GF on FLWIN and FLWOUT to smoothen the signals. Then, peaks and troughs are located by sliding windows, which perform a simple comparison to neighboring timesteps. Peaks and troughs were required to be greater than $1.5 \cdot \sigma$ within its window, where σ is the standard deviation. This family of features assists the models in learning the difference between LCI signals and pumps-on/off patterns associated with normal drilling rig activities, critical for avoiding false positive alarms.

2) WINDOW NORMALIZATION

As discussed before, the sensors and calibration frequency are different depending on the drilling rig performing the drilling operations. Additionally, drilling parameters are not limited to specific ranges of values, rather they differ based on the condition of sensors, rock being drilling, hole size, among others. As a result, the distribution of the drilling parameters data (Table 1) considerably varies from sample to sample, causing the well-known dataset and covariance shift that limits the generalization capabilities of any data-driven model [38], [39]. To avoid such generalization problems, the training data would have to include all possible ranges to avoid extrapolation, which is practically impossible as operations differ and sporadic readings are frequently present. Consequently, normalizing these data enables the models to focus on the trends rather than on the values. Magana-Mora *et al.* demonstrated that the generalization capabilities of the model considerably improved after normalizing the drilling data [8].

Normalization is also known as the ''tricks and trade'' of developing ML/DL models as it expedites training where standardization limits oscillation on the ML/DL loss function before reaching an optimal point [40]–[42]. The popular form of normalization is to simply scale input based on the mean and standard deviation of the training data. However, this approach is inappropriate when dealing with time-series data that is collected from multiple data sources (asynchronous data with inconsistent frequency) and inaccurate readings (noisy, local outliers, and mis-calibrated). Meanwhile, there are other forms of normalization, which are commonly used in computer vision and natural language processing, e.g., batch normalization [43], layer normalization [44], group normalization [45], and instance normalization [46].

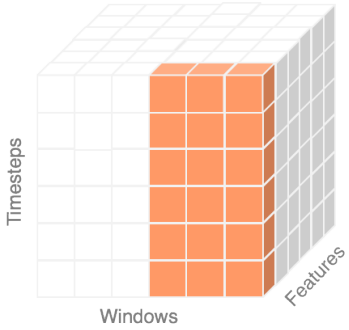

FIGURE 6. Demonstration of window normalization across grouped windows, timesteps, and features extracted from a single well. The orange columns represent an example of the scaling range, which spans all timesteps and features of training windows coming from a specific well.

In this time-series ML/DL model development, we introduce a new technique referred to as ''window normalization'' for scaling time-series data that suffer from the aforementioned quality issues. Window normalization is closest to group normalization conceptually, yet it involves practical variances. As shown in Fig. 6, window normalization does not rely on scaling statistics computed using the entire training dataset, instead, it groups windows based on wells and scales them accordingly using each batch's scaling min-max statistics. This technique aims to eliminate or minimize variations across operations (drilling rig crew practices and sensory equipment, size of the wellbore, rock heterogeneity at drilling well location, etc.). This normalization method considerably improved the performance of the models for the early detection of LCIs.

C. DATA SPLIT AND TUNING OF ML/DL ALGORITHMS

A proper data split technique is essential for avoiding model over training and optimistic validation/test results. For instance, in the extracted time-series data for LCI, adjacent samples extracted from the selected segment using a sliding window, shown in Fig. 5, have similar surface drilling parameters and sensor readings. Therefore, a random data partition to select training, validation, and test sets would induce some data leakage and would produce over-optimistic results. For instance, consider two 15-minute windows extracted at time steps t_i and t_{i+1} used for training and testing, respectively. In such a scenario, it would be almost equivalent to testing the model with training data as the testing window t_{i+1} is similar to the training sample t_i. Consequently, we divided the training, validation, and testing data at the level of the wells. More specifically, we used 80%, 10%, and 10% of the wells for training, validation, and testing, respectively. Note that for each well 60 and 180 samples are extracted for LCIs and normal operations, respectively, as seen in Fig. 5. Hence,

the number of training, validation, and testing samples (2D matrices) are 38,400 (from 160 wells), 4,800 (20 wells), and 4,800 (20 wells), respectively.

These samples are then used to derive a supervised binary classification model. However, there are many different ML algorithms that can be used to derive these models for the early detection of LCI. These include logistic regression, naïve Bayes, support vector machines, decision trees, ensemble learning, amongst others [47]. Shallow ML algorithms are generally simpler and easier to train and optimize in comparison to DL algorithms, so they cater a baseline model to evaluate more complex algorithms. Hence, we selected an ensemble learning algorithm, i.e., random forest (RF), to compare and demonstrate the effects of the window normalization scheme proposed in this study.

RF is a tree-based ML algorithm that randomly splits training samples and features into subgroups and trains multiple decision trees on each subgroup which are then combined to output a probability of the most likely class based on vote count [48]. RF is a robust (it does not require exhaustive hyperparameter tuning) and versatile model as it can handle various types of features, e.g., continuous, binary, categorical, among others. [49]. Additionally, extensive empirical studies have shown that, on average, RF outperformed other shallow learning models [50], [51]. In this study, we tuned three RF hyperparameters based on the validation set: 1) ''n_estimators'' (number of tree estimators), 2) ''max_features'' (number of features to be considered when making each tree split), and 3) ''min_sample_split'' (minimum number of samples required to split an internal node).

Meanwhile, DL algorithms provide an extended level of complexity by using deep artificial networks consisting of layers and nodes associated with nonlinear activation functions [52]. These layers and nodes interact algebraically to construct a representation that correlates input features to predict an output variable. The most basic form of these algorithms is a feedforward neural network (FNN) [53], [54]. FNN only passes node signals from one layer to another without cyclically passing data within each layer.

The disadvantage of the aforementioned ML algorithms and FNN is that they only accept a one-dimensional vector input, hence the extracted windows must be flattened out. However, flattening the input matrix results in a significantly large number of temporal features and subsequently hinders the learning process. Thus, we investigated the convolutional neural network (CNN) [55], [56] and recurrent neural network (RNN) [57], [58] algorithms, which both naturally accept a two-dimensional input matrix and directly enable their parameters to draw temporal correlations amongst drilling features. CNN and RNN architectures are most popular in computer vision and natural language processing applications, respectively. We investigated variations of both architectures in achieving early detection of LCI. Note that CNN and RNN architectures are followed by an FNN layer to reduce the output feature into the target number of classes—binary classification in this LCI application.

TABLE 3. DL hyperparametr ranges.

Hyperparameter	Tuning Range
Number of epochs	[10, 15, **20**]
Batch size	[128, **256**, 512, 1024, 2058]
Learning rate	[**0.01**, 0.001, 0.0001]
Optimizer	[SGD, Adam, **RMSprop**, Adadelta, Adagrad, Nadam]
Dropout rate	[**0.0**, 0.05, 0.1, 0.2, 0.25, 0.3]
Loss function	[class-weighted binary cross-entropy, **focal loss**]
Alpha (focal loss)	[0.3, 0.4, 0.5, 0.6, 0.7, **0.8**, 0.9]
Gamma (focal loss)	[0, **2**, 3, 4, 5]
FNN layer nodes	[16, 32, 64, **128**, 256]
RNN layer nodes	[16, 32, 64, 128, 256]
CNN layer kernel layer 1	[4, 8, **16**, 32, 64]
CNN layer kernel layer 2	[**4**, 8, 16, 32, 64]
CNN layer filter layer 1	[**8**, 16, 32, 64, 128]
CNN layer filter layer 2	[8, 16, 32, **64**, 128]
Activation function	[sigmoid, **tanh**, softmax, relu]

Note: Parameters in bold indicate the optimized values found by using a random search algorithm for the best performing CNN model.

The complexity of these models could often result in overfitting the training set [59]. Besides experimenting with different DL structures, we also applied dropout as a form of regularization [60]. The structuring and training of DL models involve tuning many hyperparameters, so we performed extensive experiments to span the various combinations of hyperparameters and allowed for various DL model parameter initializations to arrive at improved optima.

The loss function is what DL models aim to minimize during the learning process. As a result of the aforementioned aggressive sampling, the training dataset suffers from class imbalance. To mitigate this problem, we investigated class-weighted cross-entropy and focal loss [61] as the choices for the loss function. Note that the focal loss is parametrized with gamma and alpha. Gamma controls the shape of the loss cross-entropy loss function, i.e., the model can focus on the harder-to-classify training examples. Alpha introduces class weight where the rare class (LCI in this context) is emphasized further during training. We used the random search technique to tune the hyperparameters of the considered ML models based on the results obtained from the validation set. Table 3. shows the range of hyperparameters we investigated in this work.

TABLE 4. Model performance comparison.

Algorithm	Validation				Test			
	Accuracy (%)	Precision (%)	Recall (%)	F_1 Score (%)	Accuracy (%)	Precision (%)	Recall (%)	F_1 Score (%)
RF (standard normalization)	78.50	28.78	5.09	8.65	80.96	55.17	25.53	34.91
ANN (standard normalization)	78.17	43.58	31.04	36.25	89.15	79.86	61.17	69.28
RF (window normalization)	80.67	74.07	5.14	9.61	88.40	98.77	47.55	59.48
ANN (window normalization)	79.39	43.64	10.59	17.04	90.74	85.82	64.36	73.56
CNN (window normalization)	82.33	61.84	38.90	46.22	**92.55**	87.34	**73.40**	**79.77**
LSTM (window normalization)	87.64	76.52	55.09	64.06	92.45	**91.49**	68.62	78.42

III. RESULTS AND DISCUSSION

The key contributions of our study are the analysis of the sensor data obtained from two hundred historical drilling operations with LCIs to derive a robust DL model. The study emphasized on the data processing, normalization techniques, and DL model training required to derive a robust model applicable for critical drilling operations in real-time while reducing false alarms.

This section demonstrates the results obtained using the proposed methods and techniques previously described. First, we compare the performances of the different normalization approaches (standard versus window normalizations), and ML and DL models (described in the Methods Section). Note that all models discussed in this section make use of the aforementioned methods: primary and specialized preprocessing, feature engineering, aggressive sampling, and hyperparameter tuning. Table 4 compares a variety of ML/DL algorithms with and without window normalization. Trained models are compared in terms of accuracy, precision, recall, and F_1 score. While validation metrics were used to tune each model, test metrics represent each model's performance in the prediction of LCI in unseen wells.

We firstly introduce a baseline model trained using RF algorithms and standard normalization, which achieves an F_1 score of 34.91% on the test dataset. Moving over to DL models, an ANN model with standard normalization achieves an F_1 score of 69.28% on the test dataset, superior to the baseline RF model. Whereas both RF and ANN require flattening the input timestep-feature two-dimensional matrix, ANN introduces more algebraic operations that better capture the underlying nonlinearity in LCI classification. Introducing window normalization, test F_1 scores of both RF and ANN significantly improved to 59.48% and 73.56%, respectively. This improvement in evaluation metrics confirms the hypothesized impact of window normalization, which scales data across operations (drilling rig crew practices and sensory equipment, size of the wellbore, rock heterogeneity at drilling well location, etc.).

Next, we investigated the performance of CNN and LSTM algorithms, where they both feed on the input timestep-feature two-dimensional matrix without flattening. The premise of these models is to capture the temporal correlations across features to improve the differentiation of LCI versus non-LCI events. Based on the random hyperparameter tuning experimentation, the best CNN and LSTM models achieve test F_1 scores of 79.77% and 78.42%, respectively. Practically, one can argue that precision is more important than recall in this application because drilling operations are less forgiving to repeated false alarms as drilling time is costly and the crew would quickly lose confidence in such a model. Overall, optimizing both prediction and recall is most and desirable.

Although CNN and LSTM models achieve similar F_1 scores, we will only focus on the analysis visualization of the CNN model as it achieved the best performing results in terms of accuracy, recall, and F1 score. Table 3 summarizes the optimal hyperparameters of this CNN model. Arguably, expanding the hyperparameter search space may result in better model performance. However, due to the computational demands, we only considered the hyperparameter ranges described in Table 3.

It is important to note that focal loss is more effective than weighted binary cross-entropy in tackling class imbalance which stems out of the aforementioned aggressive sampling. At inference, the resultant LCI occurrence probability is to be used by the rig crew as a soft indicator. To generate a hard indicator (yes/no indicator) for LCI, then a threshold must be chosen between 0.0 and 1.0. Tuning this inference threshold on the validation F_1 score, we arrived at a threshold of 0.78.

Example #1, seen in Fig. 7, shows a scenario where severe LCI happened suddenly and the drilling rig crew reacted immediately by shutting off the pumps. The model successfully predicted this case with a probability of over 95%. Example #2, seen in Fig. 8, shows a scenario of gradual LCI, which does not result in zero FLWOUT suddenly. The drilling rig crew managed to contain the losses in the first few hours

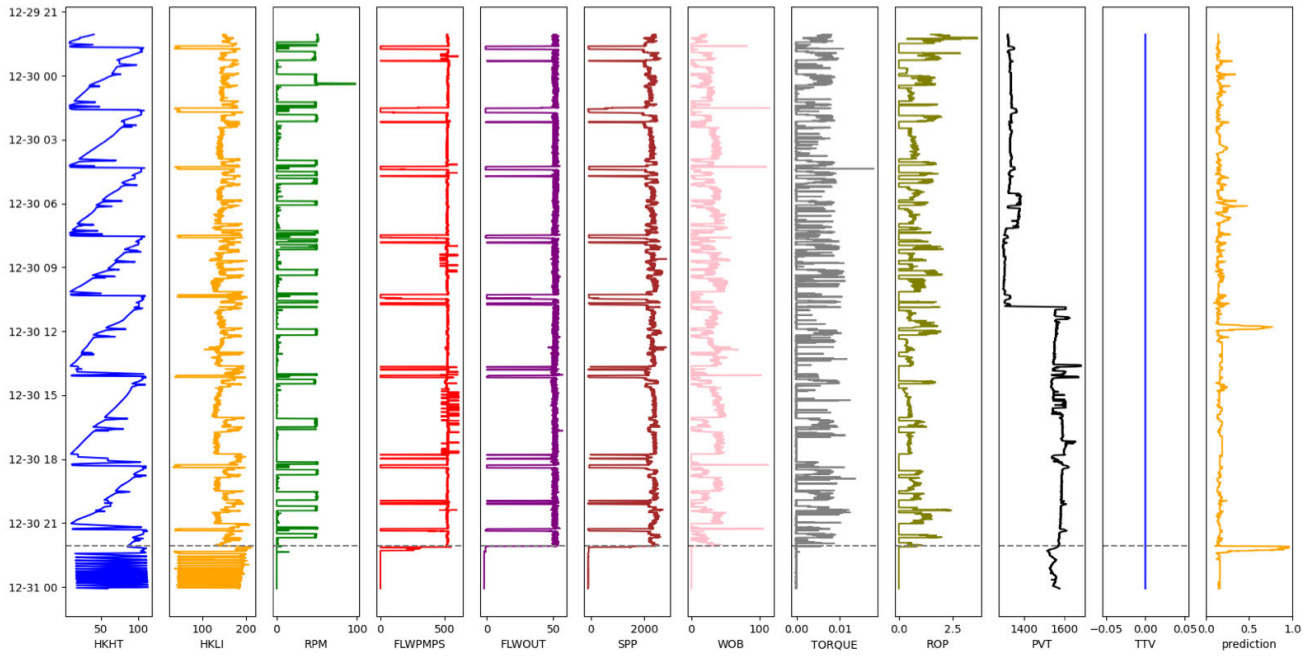

FIGURE 7. Example #1 of the optimal CNN model real-time prediction results (right column). The grey horizontal dashed line indicates when the drilling rig crew reported LCI, which reflects human performance but it does not necessarily indicate the onset of LCI. This case indicates a scenario of total losses where the drilling rig crew shut off the pumps immediately. The model successfully predicted the onset of this incident with probability of more than 95%.

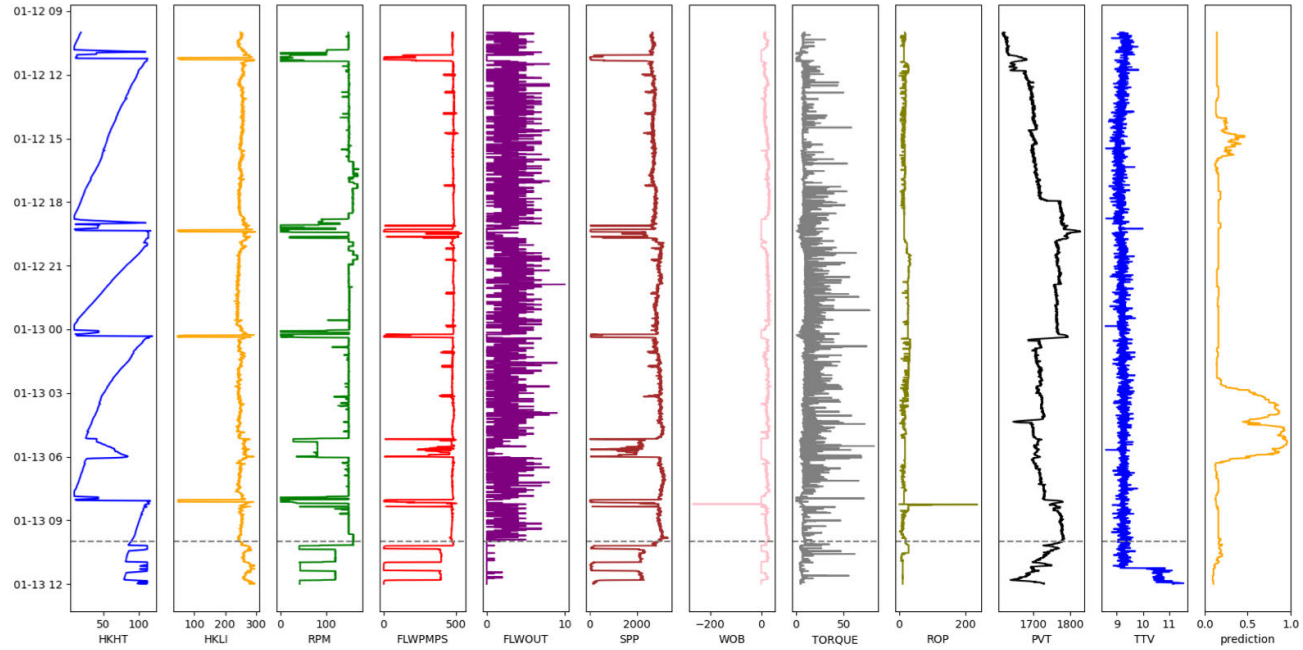

FIGURE 8. Example #2 of the optimal CNN model real-time prediction results (right column). The grey horizontal dashed line indicates when the drilling rig crew reported LCI which reflects human performance but it does not necessarily indicate the onset of LCI. This case indicates a scenario of total losses where the drilling rig crew shut off the pumps immediately. In this case, the model predicted the onset of this incident with probability of more than 90% before it was even detected by the drilling rig crew.

before they decided to shut off the pumps and attempt to cure the losses with more aggressive lost circulation material. The proposed CNN model successfully detected LCI early when it first happened, hence it is successful in the early detection of gradual LCI as well.

IV. CONCLUSION

This study demonstrated the use of advanced time-series analytics and ML/DL algorithms for the early detection of LCIs, which represent common and significant sources of NPT in drilling operations. While several models have been

proposed for the early detection of LCIs, the models have failed to properly account for the different challenges in the data quality to derive a model able to generalize to unseen drilling operations (usually performed by different drilling contractors/service providers with different sensors and calibrations). To account for these challenges, the proposed approach involves the development and implementation of a data extraction pipeline, primary and secondary preprocessing, aggressive sampling, time-series feature engineering, window normalization, and DL modeling. After exhaustive hyperparameter tuning and model exploration, results showed that CNN models performed best in the early detection of LCIs with state-of-the-art precision, recall, and F_1 scores of 87.34%, 73.40%, and 79.77%, respectively, on wells with LCIs unseen to the model (test set). Although the model was derived to detect severe/total fluid losses, the model was able to detect sign leading to seepage or partial losses. Consequently, the statistical measures used to assess the model may underestimate the model performance by penalizing correct predictions for less severe losses.

The objective of an accurate model for the detection of LCIs in real-time is to immediately detect fluid losses (seepage, partial, or complete) to enable the crew to take prompt corrective measures. For instance, seepage losses may occur due to induced fractures, which can be corrected by adjusting the drilling mud/fluid weight. Conversely, seepage losses may be observable before severe/complete losses when drilling through a cavernous formation where smaller fractures may surround a larger fracture or cavern.

Note that model development and evaluation were conducted to simulate real-time settings to fully resemble practical implications. Additionally, the model can be easily re-trained with data captured by novel sensors, such as a tuning fork that measures the drilling fluid viscosity and density in real-time [62].

Finally, the proposed approach can be applied similarly for the detection and prediction of other hazardous drilling events (stuck pipe incidents, kick/formation influx, amongst others).

REFERENCES

[1] J. Cook, F. Growcock, Q. Guo, M. Hodder, and E. van Oort, "Stabilizing the wellbore to prevent lost circulation," *Oilfield Rev.*, vol. 23, no. 4, pp. 26–35, 2011.

[2] C. Carpenter, "Liner-drilling technology mitigates lost circulation offshore mexico," *J. Petroleum Technol.*, vol. 66, no. 06, pp. 104–107, Jun. 2014.

[3] H. Wang, B. F. Towler, and M. Y. Soliman, "Fractured wellbore stress analysis: Sealing cracks to strengthen a wellbore," in *Proc. All Days*, Feb. 2007, Paper SPE-104947-MS, doi: 10.2118/104947-MS.

[4] C. P. Gooneratne, E. S. Gomez Gonzalez, A. S. Al-Musa, and H. F. Osorio, "Thirsty reservoirs-challenges in drilling through severe lost circulation zones," in *Proc. Abu Dhabi Int. Petroleum Exhib. Conf.*, 2017, Paper SPE-188461-MS, doi: 10.2118/188461-MS.

[5] H. Wang, B. F. Towler, and M. Y. Soliman, "Near wellbore stress analysis and wellbore strengthening for drilling depleted formations," in *Proc. Rocky Mountain Oil Gas Technol. Symp.*, 2007, Paper SPE-102719-MS, doi: 10.2118/102719-MS.

[6] A. Lavrov, *Lost Circulation: Mechanisms and Solutions*. Houston, TX, USA: Gulf Professional Publishing, 2016.

[7] A. Alshaikh, A. Magana-Mora, S. A. Gharbi, and A. Al-Yami, "Machine learning for detecting stuck pipe incidents: Data analytics and models evaluation," in *Proc. Int. Petroleum Technol. Conf.*, 2019, Paper IPTC-19394-MS, doi: 10.2523/IPTC-19394-MS.

[8] A. Magana-Mora, S. Gharbi, A. Alshaikh, and A. Al-Yami, "AccuPipePred: A framework for the accurate and early detection of stuck pipe for real-time drilling operations," in *Proc. SPE Middle East Oil Gas Show Conf.*, 2019, Paper SPE-194980-MS, doi: 10.2118/194980-MS.

[9] E. Davidson, L. Richardson, and S. Zoller, "Control of lost circulation in fractured limestone reservoirs," in *Proc. IADC/SPE Asia Pacific Drilling Technol.*, 2000, Paper SPE-62734-MS, doi: 10.2118/62734-MS.

[10] S. Masi, C. Molaschi, F. Zausa, and J. Michelez, "Managing circulation losses in a harsh drilling environment: Conventional solution vs. CHCD through a risk assessment," *SPE Drilling Completion*, vol. 26, no. 02, pp. 198–207, Jun. 2011.

[11] S. Wang, Y. Jiang, C. J. Zheng, B. Wu, X. Zheng, J. Tang, J. Yu, and H. Fan, "Real-time downhole monitoring and logging reduced mud loss drastically for high-pressure gas wells in tarim basin, China," *SPE Drilling Completion*, vol. 25, no. 02, pp. 187–192, Jun. 2010.

[12] F. Labenski, P. Reid, and H. Santos, "Drilling fluids approaches for control of wellbore instability in fractured formations," in *Proc. SPE/IADC Middle East Drilling Technol. Conf. Exhib.*, 2003, Paper SPE-85304-MS, doi: 10.2118/85304-MS.

[13] Y. Feng, J. F. Jones, and K. Gray, "A review on fracture-initiation and-propagation pressures for lost circulation and wellbore strengthening," *SPE Drilling Completion*, vol. 31, no. 02, pp. 134–144, 2016.

[14] Y. Feng and K. E. Gray, "Review of fundamental studies on lost circulation and wellbore strengthening," *J. Petroleum Sci. Eng.*, vol. 152, pp. 511–522, Apr. 2017.

[15] S. Salehi and R. Nygaard, "Numerical modeling of induced fracture propagation: A novel approach for lost circulation materials (LCM) design in borehole strengthening applications of deep offshore drilling," in *Proc. SPE Annu. Tech. Conf. Exhib.*, 2012, Paper SPE-135155-MS, doi: 10.2118/135155-MS.

[16] J. Ramasamy and M. Amanullah, "Two component lost circulation material for controlling seepage to moderate losses," in *Proc. SPE Kingdom Saudi Arabia Annu. Tech. Symp. Exhib.*, 2017, Paper SPE-188101-MS, doi: 10.2118/188101-MS.

[17] J. Ramasamy, M. Amanullah, and M. K. Al-Arfaj, "A large aperture test apparatus for severe lost circulation material evaluation," in *Proc. SPE Kingdom Saudi Arabia Annu. Tech. Symp. Exhib.*, 2018, pp. 1–4.

[18] J. Ramasamy, C. Gooneratne P, and M. Amanullah, "Current methods and novel solutions for mitigating lost circulation," in *Proc. Int. Petroleum Technol. Conf.*, 2019, Paper IPTC-19499-MS, doi: 10.2523/IPTC-19499-MS.

[19] R. Ahmed and S. Miska, "Advanced wellbore hydraulics," in *Advanced Drilling and Well Technology: Society of Petroleum Engineers*. Richardson, TX, USA: Society of Petroleum Engineers, 2009, pp. 191–220.

[20] M. Bassam, "Effect of cuttings on annular pressure loss," M.S. thesis, Dept. Petroleum Geosyst. Eng., Univ. Texas, Austin, TX, USA, 2020.

[21] Z. Li, M. Chen, Y. Jin, Y. Lu, H. Wang, and Z. Geng, "Study on intelligent prediction for risk level of lost circulation while drilling based on machine learning," in *Proc. 52nd US Rock Mech./Geomech. Symp.*, 2018, Paper ARMA-2018-105.

[22] M. Sabah, M. Mehrad, S. B. Ashrafi, D. A. Wood, and S. Fathi, "Hybrid machine learning algorithms to enhance lost-circulation prediction and management in the marun oil field," *J. Petroleum Sci. Eng.*, vol. 198, Mar. 2021, Art. no. 108125.

[23] A. K. Abbas, A. A. Bashikh, H. Abbas, and H. Q. Mohammed, "Intelligent decisions to stop or mitigate lost circulation based on machine learning," *Energy*, vol. 183, pp. 1104–1113, Sep. 2019.

[24] A. T. T. Al-Hameedi, H. H. Alkinani, S. Dunn-Norman, R. E. Flori, S. A. Hilgedick, A. S. Amer, and M. Alsaba, "Mud loss estimation using machine learning approach," *J. Petroleum Explor. Prod. Technol.*, vol. 9, no. 2, pp. 1339–1354, Jun. 2019.

[25] A. Moazzeni, M. Nabaei, and S. G. Jegarluei, "Prediction of lost circulation using virtual intelligence in one of Iranian oilfields," in *Proc. Nigeria Annu. Int. Conf. Exhib.*, 2010, Paper SPE-136992-MS, doi: 10.2118/136992-MS.

[26] A. Ahmed, S. Elkatatny, A. Ali, and A. Abdulraheem, "Prediction of lost circulation zones using artificial neural network and functional network," in *Proc. Abu Dhabi Int. Petroleum Exhib. Conf.*, 2020, Paper SPE-203268-MS, doi: 10.2118/203268-MS.

[27] A. Magana-Mora, M. Abughaban, and A. Ali, "Machine-learning model for the prediction of lithology porosity from surface drilling parameters," in *Proc. Abu Dhabi Int. Petroleum Exhib. Conf.*, 2020, Paper SPE-203213-MS, doi: 10.2118/203213-MS.
[28] C. P. Gooneratne, A. Magana-Mora, W. C. Otalvora, and M. Affleck, "Drilling in the fourth industrial revolution—Vision and challenges," *IEEE Eng. Manage. Rev.*, vol. 48, no. 4, pp. 144–159, 4th Quart., 2020.
[29] C. P. Gooneratne, A. Magana-Mora, M. Affleck, W. C. Otalvora, G. D. Zhan, and T. E. Moellendick, "Camera-based edge analytics for drilling optimization," in *Proc. IEEE Int. Conf. Edge Comput. (EDGE)*, Oct. 2020, pp. 111–115.
[30] M. Khudiri, J. James, M. Amer, B. Otaibi, M. Nefai, and J. Curtis, "The integration of drilling sensor real-time data with drilling reporting data at Saudi Aramco using WITSML," in *Proc. SPE Intell. Energy Conf. Exhib.*, 2014, Paper SPE-167873-MS, doi: 10.2118/167873-MS.
[31] S. C. H. Geekiyanage, A. Tunkiel, and D. Sui, "Drilling data quality improvement and information extraction with case studies," *J. Petroleum Explor. Prod. Technol.*, vol. 11, no. 2, pp. 1–19, Nov. 2020.
[32] G. Dong and H. Liu, *Feature Engineering for Machine Learning and Data Analytics*. Boca Raton, FL, USA: CRC Press, 2018.
[33] A. Zheng, A. Casari, *Feature Engineering for Machine Learning: Principles and Techniques for Data Scientists*. Sebastopol, CA, USA: O'Reilly Media, 2018.
[34] M. Kalkatawi, A. Magana-Mora, B. Jankovic, and V. B. Bajic, "DeepGSR: An optimized deep-learning structure for the recognition of genomic signals and regions," *Bioinformatics*, vol. 35, no. 7, pp. 1125–1132, Apr. 2019.
[35] S. Hansun, "A new approach of moving average method in time series analysis," in *Proc. Conf. New Media Stud. (CoNMedia)*, Nov. 2013, pp. 1–4.
[36] R. A. Haddad and A. N. Akansu, "A class of fast Gaussian binomial filters for speech and image processing," *IEEE Trans. Signal Process.*, vol. 39, no. 3, pp. 723–727, Mar. 1991.
[37] M. Nixon and A. Aguado, *Feature Extraction and Image Processing for Computer Vision*. New York, NY, USA: Academic, 2019.
[38] J. G. Moreno-Torres, T. Raeder, R. Alaiz-Rodríguez, N. V. Chawla, and F. Herrera, "A unifying view on dataset shift in classification," *Pattern Recognit.*, vol. 45, no. 1, pp. 521–530, Jan. 2012.
[39] J. Quiñonero-Candela, M. Sugiyama, N. D. Lawrence, and A. Schwaighofer, *Dataset Shift in Machine Learning*. Cambridge, MA, USA: MIT Press, 2009.
[40] M. J. Aljubran and R. Horne, "Prediction of multilateral inflow control valve flow performance using machine learning," in *Proc. SPE Annu. Tech. Conf. Exhib.*, 2019.
[41] G. Joel, *Data Science From Scratch*. Sebastopol, CA, USA: O'Reilly Media, 2015.
[42] Y. A. LeCun, L. Bottou, G. B. Orr, and K.-R. Müller, "Efficient backProp BT-neural networks: Tricks of the trade," in *Neural Networks: Tricks of the Trade*. Berlin, Germany: Springer, 2012.
[43] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 448–456.
[44] J. Lei Ba, J. Ryan Kiros, and G. E. Hinton, "Layer normalization," 2016, *arXiv:1607.06450*. [Online]. Available: http://arxiv.org/abs/1607.06450
[45] Y. Wu and K. He, "Group normalization," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 3–19.
[46] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Instance normalization: The missing ingredient for fast stylization," 2016, *arXiv:1607.08022*. [Online]. Available: http://arxiv.org/abs/1607.08022
[47] G. Bonaccorso, *Machine Learning Algorithms*. Birmingham, U.K.: Packt Publishing, 2017.
[48] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
[49] A. Cutler, D. R. Cutler, and J. R. Stevens, "Random forests," in *Ensemble Machine Learning*. Boston, MA, USA: Springer, 2012, pp. 157–175.
[50] A. Magana-Mora and V. B. Bajic, "OmniGA: Optimized omnivariate decision trees for generalizable classification models," *Sci. Rep.*, vol. 7, no. 1, pp. 1–11, Dec. 2017.
[51] M. Fernández-Delgado, E. Cernadas, S. Barro, and D. Amorim, "Do we need hundreds of classifiers to solve real world classification problems?" *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 3133–3181, 2014.
[52] T. M. Mitchell, "Artificial neural networks," *Mach. Learn.*, vol. 45, pp. 81–127, Mar. 1997.
[53] A. Zell, *Simulation Neuronaler Netze*, nos. 5–3. Reading, MA, USA: Addison-Wesley, 1994.
[54] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015.
[55] K. O'Shea and R. Nash, "An introduction to convolutional neural networks," 2015, *arXiv:1511.08458*. [Online]. Available: http://arxiv.org/abs/1511.08458
[56] J. Gu, Z. Wang, J. Kuen, L. Ma, A. Shahroudy, and B. Shuai, "Recent advances in convolutional neural networks," *Pattern Recognit.*, vol. 77, pp. 354–377, May 2018.
[57] L. R. Medsker and L. Jain, "Recurrent neural networks," *Des. Appl.*, vol. 5, pp. 2–3, Dec. 2001.
[58] R. Pascanu, C. Gulcehre, K. Cho, and Y. Bengio, "How to construct deep recurrent neural networks," 2013, *arXiv:1312.6026*. [Online]. Available: http://arxiv.org/abs/1312.6026
[59] T. Dietterich, "Overfitting and undercomputing in machine learning," *ACM Comput. Surveys*, vol. 27, no. 3, pp. 326–327, Sep. 1995.
[60] P. Baldi and P. J. Sadowski, "Understanding dropout," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 26, 2013, pp. 2814–2822.
[61] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, "Focal loss for dense object detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Oct. 2017, pp. 2980–2988.
[62] M. Gonzalez, T. Thiel, C. Gooneratne, R. Adams, C. Powell, A. Magana-Mora, J. Ramasamy, and M. Deffenbaugh, "Development of an in-tank tuning fork resonator for automated viscosity/density measurements of drilling fluids," *IEEE Access*, vol. 9, pp. 25703–25715, 2021.

MOHAMMAD ALJUBRAN received the M.S. degree in energy resources engineering from Stanford University, CA, USA. During his M.S. studies, he developed advanced deep learning models for various energy applications, including wind turbine detection and characterization using satellite imagery, optimization and uncertainty quantification of oil and gas resources, and webpage with live natural language analyzers of the geothermal literature. He currently works as a Data Scientist with the EXPEC Advanced Research Center, Saudi Aramco, where he is researching, developing, and drilling hazard prediction and optimization tools using deep learning algorithms. He has published multiple peer-reviewed articles and authored several granted patents.

JOTHIBASU RAMASAMY received the Ph.D. degree in chemistry from the National University of Singapore (NUS), Singapore. During his Ph.D. and post-doctoral fellowships at NUS and the King Abdullah University of Science and Technology, Saudi Arabia, he worked on developing novel organometallic compounds for homogeneous catalysis that resulted in several peer-reviewed journal publications and conference papers. He currently works as a Petroleum Scientist with the EXPEC Advanced Research Center (EXPEC ARC), Saudi Aramco, where he works on developing novel additives for drilling fluids. His research interests include developing new chemicals by recycling used materials and environment-friendly drilling fluid additives.

MOHAMMED ALBASSAM received the master's degree in petroleum engineering from The University of Texas at Austin, in 2020. In 2015, he joined Saudi Aramco as part of the Upstream Research Center (EXPEC ARC), where he currently works as a Petroleum Engineer, developing new techniques to calculate annular pressure loss while drilling. His research interest includes establishing connection between bottom-hole and surface to acquire downhole data and measurements.

ARTURO MAGANA-MORA received the Ph.D. degree in computer science from the King Abdullah University of Science and Technology, Saudi Arabia. During his Ph.D. studies and a Postdoctoral Fellowship at the National Institute of Technology (AIST), Japan, he developed novel artificial intelligence models to address problems in biology, genomics, and chemistry that resulted in several peer-reviewed publications in high-impact journals, poster presentations, and invited talks. He is currently the Lead Data Scientist with the Drilling Technology Division, EXPEC Advanced Research Center, Saudi Aramco, where he has unveiled new opportunities in the domain of drilling automation and optimization, and catalyzed existing work. During his career, he has used his expertise in computer science to bridge artificial intelligence with biology, genomics, chemistry, and the oil and gas industry. He also serves as an associate editor and a referee for several scientific journals.