Stuck Pipe Prediction in Deep Wellbores Drilled in Complex Evaporite Formations using ML-Based Intelligent Classifiers

By

Elmira Amanzhol

Thesis submitted to the School of Mining and Geosciences of Nazarbayev University in Partial Fulfillment of the Requirements for the Degree of

Master of Science in Petroleum Engineering

Nazarbayev University

April 2025

ORIGINALITY STATEMENT

I, Elmira Amanzhol, hereby declare that this submission is my own work and to the best of my knowledge it contains no materials previously published or written by another person, or substantial proportions of material which have been accepted for the award of any other degree or diploma at Nazarbayev University or any other educational institution, except where due acknowledgement is made in the thesis.

Any contribution made to the research by others, with whom I have worked at NU or elsewhere is explicitly acknowledged in the thesis.

I also declare that the intellectual content of this thesis is the product of my own work, except to the extent that assistance from others in the project's design and conception or in style, presentation and linguistic expression is acknowledged.

Signed on **15.04.2025**

____

ABSTRACT

Stuck pipe events represent a very challenging and high-cost problem in drilling industry. Hence, accurate stuck pipe prediction is crucial for successful field development. Various stuck pipe prediction models using Machine Learning (ML) approaches have been developed in the past. In this study, an attempt was made to address critical problem of stuck pipe incidents prediction while drilling through deep and complex evaporite formations by ML-based models for early hazard detection. Development and testing different intelligent models performed using real field data that included actual drilling parameters along with geological information and drilling mud properties. Actual field data from 10 wells were used to train and test the models using acquired 61 data sets that consist of 610 datapoints. Five supervised classification ML algorithms were used: Logistic Regression (LR), K-Nearest Neighbors (KNN), Support Vector Machines (SVM), Decision Trees (DT) and CatBoost. Before proceeding with training phase, an essential step of data pre-processing was conducted to identify any missing data and detecting outliers. Each model was trained and tested using the 60-40 data splitting strategy to identify stuck pipe condition at both individual hole section level and the entire field level. The developed intelligent models can identify patterns of different stuck pipe conditions including stuck pipe, non-stuck, pack-off, overpull, and tight spot. Error metrices were used to ascertain accuracy, precision, recall, and F1-score of the developed intelligent models in conjunction with ROC analysis to assess performance of the developed intelligent models. Two intelligent models were identified as the most effective classifiers for specified stuck pipe issues: DT model attained 99.59% accuracy at the field level; while, CatBoost reached 100% accuracy in hole section-based assessments. Feature importance analysis together with SHAP (SHapley Additive exPlanations) analysis showed that formation lithology is the leading factor affecting stuck pipe occurrences, as salt and clay contents remain as a primary contributors to this issue. Two key drilling parameters that control stuck pipe occurrence were identified as drilling rotational speed and flow rate. Additionally, mud weight combined with gel strength demonstrated major effects of rheological properties of the mud. SHAP analysis showed that some hole sections outcomes are locally affected by depth and clay content levels. The results of this research demonstrated that successful field development operations require interpretable intelligent models to enhance operational decision-making process to prevent different drilling hazards.

Keywords: stuck pipe, stuck pipe prediction, evaporites, salt, mechanical sticking, differential sticking, machine learning (ML), intelligent classifiers

ACKNOWLEDGMENTS

I want to extend my sincere appreciation to my supervisor, Dr. Ali Shafiei, who provided essential guidance and expert direction with unfailing support during my research process. His feedback and encouragement were indispensable factors which helped me to develop this research work.

I would like to appreciate the thesis committee members Dr. Sohrab Zendehboudi and Dr. Shams Kalam for constructive feedback, critical evaluation and suggestions which enhanced the quality of this research. I express my deep gratitude to Ghazal Piroozi, Research Assistant, who provided essential support and assistant while working on AI-based modeling and results interpretation.

I also would like to express my heartfelt gratitude to my family, who persistently believed in me and continuously encouraged throughout the time I dedicated to this project. The thesis presents my efforts together with important inputs from all contributors mentioned in this study.

Table of Contents

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

A brief overview of drilling operations, stuck pipe conditions while drilling, primary mechanisms proposed for pipe sticking, its repercussions to drilling industry, and remedy actions related to release the stuck pipe are presented in Chapter 1. The section is followed by problem statement, relevance of the project to the industry, objectives of the research, and main machine learning (ML) algorithms used in this research work. The chapter concludes with presentation of the thesis structure.

1.1 Background

Drilling a wellbore is known as the costliest upstream operation in oil and gas industry (Albaiyat, 2012). Constructing a wellbore is a complex process that includes breaking the rock, removal of rock cuttings from the hole, and strengthening the wellbore through casing and cementing activities, depending on objectives of wellbore construction operation, reservoir geology, and wellbore trajectory. Drilling a well is usually followed by consecutive actions of drilling the rock, reaming operations to clean and maintain the well bore stability, tripping in and out of the hole to change the bottom hole assembly (BHA), connecting the new components to the drill string, circulations to clean up the well bore (Brankovic et al., 2021). Drill string typically consists of BHA required to drill the well, drill bit used to break the rock and drill pipes. Depending on the well objectives and its complexity, BHA could consist of advanced technological equipment required to drill the well such as measurements while drilling (MWD) and logging while drilling (LWD) tools, downhole mud motors, rotary steerable systems (RSS), hole enlargement equipment (reamer), and coring assembly. Drilling activities result in a number of challenges such as stuck pipe situation, lost circulation, wellbore stability issues, fluid influx, problems related to well control (Shahbazi & Shahri, 2011).

Among all of these drilling problems, stuck pipe incidents raise the drilling industry's expenses the most. This is because they lead to losing the objectives of wellbore construction, significant Non-Productive Time (NPT) for remedial activities to release the stuck pipe and re-drill the well. According to statistics reported in the literature, stuck pipe incidents account for 25% of NPT out of all drilling issues (Muqeem et al., 2012), in some places this number can reach up to 52% (Jardine et al., 1992). Furthermore, it might result in substantial losses because of expensive downhole equipment (e.g., LWD) damage or loss of the wellbore. As part of LWD tools, some downhole assemblies may contain radioactive (RA) sources, losing the RA source downhole brings additional environmental consequences that can lead to formation contamination and regulatory problems for the companies. Occurrence of stuck pipe accidents varies greatly depending on type of operation; around 50% of all incidents correspond to tripping in or out of the hole, 20% occurs while going for connection the drill pipe and reaming activities, and 10% of stuck events take place while on bottom drilling (Bradley et al., 1991).

Stuck pipe condition is known as a situation when the drill string loses its ability to move in any direction and become unable to rotate the string (Tsuchihashi et al., 2021). In some conditions, no circulation through the drill string is possible, which worsen the situation. There are several mechanisms that causes the pipe to stick, which highly dependent on geology of formation, pore pressure, well trajectory, wellbore cleaning issues, incompatible drilling fluids, bottom hole assembly design (Elmousalami & Elaskary, 2020). There are two main stuck pipe categories defined in the oil and gas industry (Murillo et al., 2009):

Mechanical stuck pipe occurs when there is a physical barrier down the hole that keeps the string from free movement. The main reasons for mechanical stuck pipe are wellbore geometry, downhole junk or wellbore instability because of accumulated cuttings that pack-offs the drill string (Issa et al., 2023).

Differential sticking occurs when there are two opposing pressure forces in wellbores which lead to differential sticking development: hydrostatic pressure inside the wellbore and the formation pressure. Drilling operations achieve overbalanced drilling through maintaining hydrostatic pressures greater than formation pressure to prevent fluid from entering the hole. The resulting force from pressure differential pushes the drill string towards the wall where it sticks to the wellbore. Usually, this type of sticking develops when drilling through permeable rock formations (Shadizadeh et al., 2010).

1.2 Problem definition

The problem of stuck pipe prediction has drawn attention of numerous scholars for decades. The approaches developed over the years in drilling industry could be classified into following categories: analytical models, data-driven approaches and hybrid methods which incorporate physics-based models with data-driven approaches (S. Zhu et al., 2022).

Because of complex structure and physics of pipe sticking state, it is extremely difficult to anticipate and identify the right mechanism at early times, allowing for prompt use of respective actions. Within the first four hours after the incident, 50% of stuck pipe cases have been released successfully. While, less than 10% of the cases were released when time was greater than 4 hours and roughly 40% of the cases were never released (Alshaikh et al., 2018). Hence, preventing any possibility of stuck pipe is at the greatest importance rather than mitigating the repercussions of such incidents (Shahbazi & Shahri, 2011). Accurate classification of pipe sticking conditions will aid in defining the most appropriate drilling strategy for future field development to avoid the risk of pipe sticking.

Furthermore, effectiveness of model predictions can vary with specific stuck pipe features influenced by mechanism of stuck pipe, location, formation lithology, and well trajectory (Abbas et al., 2019). As an example, 69.5% of all stuck incidents of Saudi Aramco wells in 2009 caused by mechanical stuck (Muqeem et al., 2012); while differential sticking was the main cause for 61% of non-productive expenses in the Gulf of Mexico (Mahmood & Assi, 2024). Hence, the intricacy of input parameters and their non-linear correlations require searching for alternative solutions for conventional statistical analysis (Abbas et al., 2019).

A number of research works demonstrated that ML-based approaches can offer accurate and reliable solution for stuck pipe prediction problems (Albaiyat, 2012; Jahanbakhshi et al., 2012; Zhao et al., 2017; Magana-Mora et al., 2019; Alshaikh et al., 2019; Abbas et al., 2019; Elmousalami & Elaskary, 2020; Brankovic et al., 2021; Payrazyan & Robinson, 2023). The intelligent models were developed based on a number of drilling parameters and showed high accuracy, > 80%. Despite the fact that ML techniques demonstrated high efficiency and accuracy in dealing with drilling data for stuck pipe prediction, the primary challenges of this approach still include data availability, uncertainty of the results, generalizability, early sign detection of stuck pipe condition, false positive alarms, and overfitting problems.

1.3 Relevance to the industry

In addition to substantial financial losses, stuck pipe events remain a leading challenge in drilling operations because of their potential safety risks and environmental consequences (D'Amicis et al., 2023). The recent statistics represents some improvements in NPT percentage caused by stuck pipe incidents that is about 15%. However, the credit for this mainly goes to preventative measures that were actively applied in drilling operations based on previous lessons learned such as appropriate hole cleaning techniques, optimized drilling parameters, and suitable BHA design (Alshaikh et al., 2019). At the same time, monitoring of drilling parameters towards automation to detect stuck pipe situation in real time also made some positive impact on prevention of stuck pipe incidents (Alshaikh et al., 2018). However, statistical methods that underlie automation approaches have serious limitations primarily related to nature of stuck occurrences and complex input parameters patterns, which renders this method inefficient in early stuck pipe prediction and results in a high number of false alarms (Abbas et al., 2019). Hence, a developing more robust and highly accurate ML-based intelligent models can notably expand this domain and provide strong instruments to create efficient and accurate stuck pipe predictions. Nevertheless, these approaches are not widely utilized yet alongside traditional methods in the drilling industry. Further research in this domain is still necessary to address main shortcomings of these approaches and to develop more robust models for more challenging formations. In addition, adoption of new ML methods within oil and gas industry continues to rise, further expanding the opportunities of evaluation of different intelligent methods.

1.4 Research objectives

The primary objective of this research was to develop and evaluate performance of different ML-based intelligent models and to identify the most effective and accurate intelligent model for stuck pipe prediction in deep complex evaporite formations using actual field data obtained from 10 wells which included 30 stuck pipe and 31 non-stuck cases. The main objectives of this research work are listed below:

To conduct a comprehensive review of prior studies on stuck pipe prediction using different intelligent models based on ML approaches, to determine the most efficient algorithms to be used in this project. To develop intelligent models using field data.

To assess performance of the different developed smart models using specified error metrices, as a way to define the best intelligent model that will demonstrate the highest accuracy and produce reliable stuck pipe prediction model for the given field.

To perform sensitivity analysis by using feature importance analysis techniques and to determine which of the input parameters are influencing the output for each hole section and the whole field for the purpose of identifying correct remedy actions and stuck-pipe free drilling strategy for future development of the field.

1.5 Research methodology

This research work aims at developing accurate and robust stuck pipe prediction intelligent models using ML approaches for a particular field to define the correct strategy for future "stuck pipe-free drilling". The research commenced from a comprehensive analysis of stuck pipe prediction techniques using ML algorithms where the main approaches were established to be used in the research, such as Logistic Regression algorithm (LR), K-nearest neighbors (KNN) algorithm, Support Vector Machine (SVM), Decision Tree and the recent CatBoost – gradient boosting method, which has not been used previously. The model development process started with collecting input data and creating the datasets, as this step is considered as the most critical for success of the model and a very time-consuming activity (Elmousalami & Elaskary, 2020). The process is followed by data quality evaluation and pre-processing tasks which involve evaluating quality of obtained data through input analysis for outlier detection and missing data identification. The acquired data was then labeled according to the state of drilling activity, hole section, and stuck pipe condition. Separate training and testing data sets from the obtained field data was used. Each selected ML algorithm was run and tested using the assembled datasets, and accuracy of each method was evaluated based on predefined error metrics. The models with the highest accuracy were then chosen for the further sensitivity analysis to identify the feature importance parameters for each hole section to define the proper course of mitigation actions and way forward for future stuck pipe-free drilling in the field.

1.6 Thesis structure

The thesis consists of five chapters. The first chapter contains the background information on stuck pipe incidents and its relevance to the oil and gas industry. It also includes the problem statement, goals and objectives of the proposed research work, and a brief description of the research methodology used. The second chapter is focused on the literature review, where the main studies that were conducted in the field of stuck pipe prediction using variety of approaches, such as analytical approaches, data-driven approaches and hybrid models are presented. The main emphasis went toward the ML approaches application within the research area. Additionally, stuck pipe condition, main mechanisms, and remedy actions to release the stuck pipe are discussed. A detailed information of different ML algorithms and step by step procedures followed to develop each model are presented in chapter three. The results of training and testing the models with comparison of evaluation metrics are presented and discussed in depth in chapter four. In addition, sensitivity analysis of input parameters is performed for selected model. At the end, the main outcomes of the research work are presented as conclusions section in chapter five along with the way forward of future projects.

2. LITERATURE REVIEW

A literature review of main drilling problems associated with stuck pipe, their key mechanisms, and prediction methods used in the industry, along with their advantages and disadvantages are presented in this chapter. The first part covers background information on the stuck pipe mechanisms and the primary challenges associated with them. This is followed by discussion of main causes and conditions for their occurrence, as well as preventative and remedial actions to release the stuck pipe. Challenges associated with drilling mobile formations such as salt formations are also discussed. Detailed information on prediction approaches used in the industry, encompassing physics-based approach, data-driven approaches, and various hybrid approaches is provided in second part of the chapter. Finally, the previously proposed intelligent prediction methods are examined and with their accompanying performance evaluation, benefits, and limitations.

2.1 Background

Drilling a well stands as an essential oil and gas industry practice that allows to locate hydrocarbons and supports continuous operations from the initial production to the final phase. Geological structure, formation composition, hydrocarbon type, engineering purpose of the well, its design and trajectory can significantly vary depending on location of the oil and gas fields. This results in numerous challenges associated with drilling process. Among the various drilling challenges, stuck pipe is considered as rare events, however with notable consequences to the entire process (Brankovic et al., 2021).

Different physical phenomena are associated with the stuck pipe condition, allowing to classify them into two main categories: mechanical stuck pipe and differential sticking. Where all categories, except for differential sticking, are classified as sub-categories of mechanical stuck pipe. Some studies define three categories of pipe sticking condition depending on the mechanisms, such as pack-off and bridging, wellbore geometry induced and differential stuck pipe (Mitchell, 2001; Meor Hashim et al., 2021).

2.1.1 Mechanical sticking

In the event of mechanical stuck pipe, the drill string loses its ability to move because of the existence of physical barrier or blockage in the borehole. The primary causes of mechanical sticking can be classified into two sub-categories: pack-off and bridging, and well geometry induced stuck pipe (Shadizadeh et al., 2010). Pack-off typically occurs because of the accumulation of small cuttings around the area of drill string restricting circulation and impeding free movement. If not addressed immediately, the issue can rapidly escalate into a stuck pipe condition. Bridging usually occurs when medium to large fragments of rock or cement clog the annulus surrounding the drill pipe (Figure 2-1). Which facilitates limited fluid flow because of the partial or complete blockage of the pipe (Willersrud, 2015). Several factors contribute to the string packing off or bridging: Drilling through unconsolidated formations, reactive formations (e.g., water sensitive shales), mobile formations (e.g., salt), insufficient hole cleaning, and ineffectiveness of drilling mud to suspend cuttings.

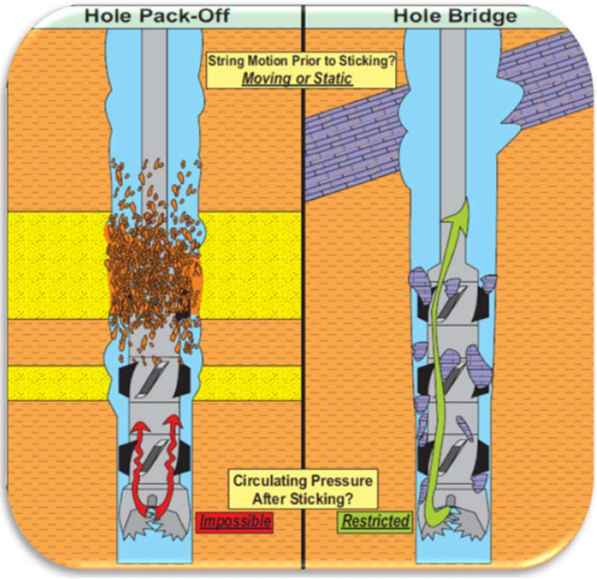
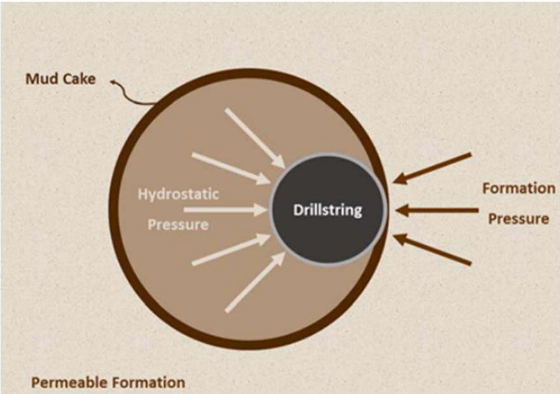


Figure 2-1 – Mechanical stuck pipe illustrations of pack-off and bridging (Drilling Manual, https://www.drillingmanual.com/)

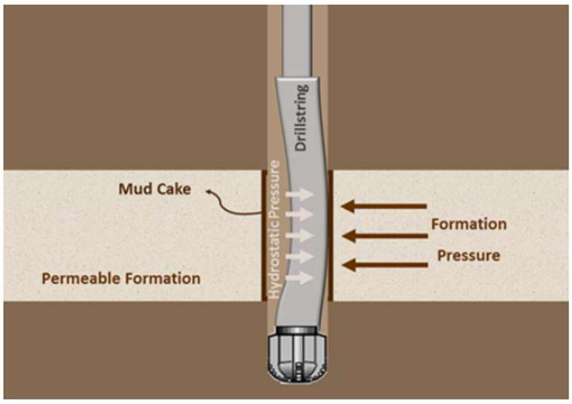
Well geometry originated stuck mechanism arises from the shape and condition of the bore hole. One reason pertains to the under-gauge hole, which is caused by hard or abrasive drilling formations, leading to considerable wear on the bit and stabilizer, resulting in the hole diameter being smaller than the actual bit size. When passing these intervals, the string may become physically caught. Additionally, key seats can induce stuck pipe situations because of well geometry, since the rotation of the drill string against the bore hole may create a groove or key seat in the wall where BHA components, such as stabilizers, might become stuck (Aljubran et al., 2017). It is induced by a sudden change in inclination or azimuth, or because of a transition from medium to hard formation. Stuck conditions associated with wellbore geometry also encompass well tortuosity issues, wherein the well trajectory deviates significantly (macro tortuosity – more than 30 m) or slightly (micro tortuosity – 1m) from the planned or smooth well pass because of formation characteristics, drilling technologies, or other aspects (D'Angelo et al., 2019). Additionally, there are several other factors contributing to mechanical stuck pipe condition (Bowes et al., 1997): High doglegs, junk or debris in the hole, cement blocks, and green cement.

2.1.2 *Differential sticking*

This category encompasses stuck pipe events caused by hydrostatic pressure of the mud column in the borehole differing from the formation pressure at the drill string and rock interface (Brandon et al., 1993). This mechanism was initially identified by Helmick and Longley in 1957 and Outmans in 1958, based on laboratory studies. Depending on location, differential sticking accounts for up to 32% of all stuck incidents around the world (Muqeem et al., 2012).



(a)



(b)

Figure 2-2 - Differential pipe sticking schematics: a) forces across the pipe; b) forces across the drill string view (Alshaikh et al., 2018).

Differential sticking typically occurs when the drill string is stationary or moves at very low speed against permeable rocks under conditions of overbalance. Another key aspect is existence of filter cake in the contact zone between drill pipe and formation (Naraghi & Jamshidi, 2013). A multiple studies and experiments have been conducted to comprehend phenomenon of differential pipe sticking and impact of filter cake on it (Helmick & Longley A. J., 1957; Outmans, 1958; Hunter et al., 1978, Courteille & Zurdo, 1985; Isambourg et al., 1999). Filter cake is typically forms in regions adjacent to permeable area, where mud filtrate invades permeable zone, depositing solid particles that create a barrier between the borehole and formation. The presence and quality of mud cake are critical factors in wellbore stability issues. And when the drill pipe remains stationary against the wellbore with formed filter cake, it may become embedded in the borehole wall because of pressure differentials, resulting in triple bond mechanism between the pipe, cake, and formation as shown in
(b)

Figure 2-2. Several experiments indicated that the force exerted by the pipe against the cake is contingent upon the cake's capability to lose its permeability under compression. During rotation of the string, dynamic filtering occurs, maintaining stable pressure inside the cake. However, when the pipe becomes stationary, the pressure across the contact area with formation decreases rapidly, indicating that static filtering is occurring (Courteille & Zurdo, 1985). As the filtrate continues to leave the cake, the force between the solid particles of the cake and the pipe intensifies. In low permeable formations, this stress usually occurs in the internal part of the cake, hence exerting reduced tension across the cake. However, in high permeable formations, this tension developed around the cake, greatly elevating the likelihood of pipe sticking. Consequently, considerable force should be applied to overcome the shear force within the cake and the contact force between the pipe and the cake as shown in Figure 2-3 (Dupriest et al., 2011). Although there is no physical barrier such as during mechanical pipe stuck, the consequences of differential sticking are nonetheless severe, potentially resulting in lost equipment in the hole necessitating plugging and re-drilling the well.

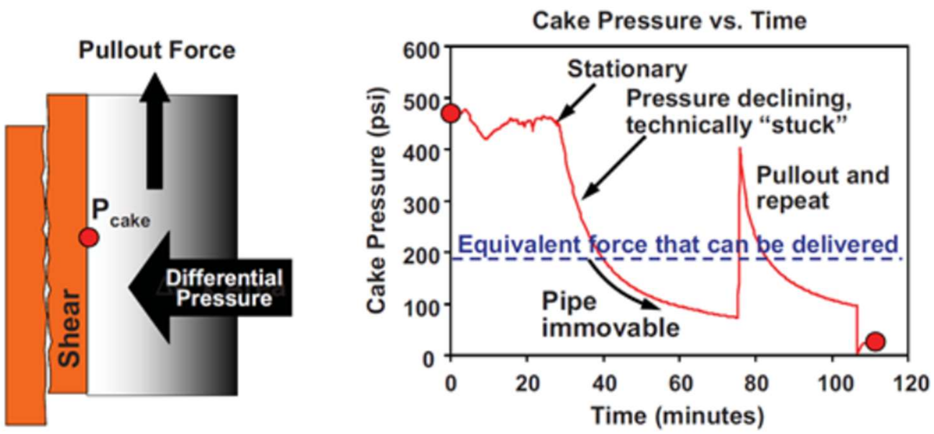


Figure 2-3 – Schematical illustration of the shear force and internal pressure of the filter cake at the different stages of the pipe motion obtained with sticking test apparatus (Dupriest et al., 2011).

2.2 Prevention and remedy measures to free the stuck pipe

As highlighted previously, the stuck pipe consequences have significant economic, environmental and safety impact to the industry. Hence, taking appropriate mitigation steps are essential to avoid catastrophic outcomes of such incidents. When dealing with stuck pipe situation, it is vital to response to the problem rapidly. The incident usually begins with some difficulties to move or rotate the string; however, it very quickly turns into complete immobility. For instance, the success rate of freeing the stuck pipe is approximately 50% within the initial four hours following the incident, and only 10% when the duration exceeds 4 hours (Mitchell, 2001). Thus, the timely execution of appropriate remedial measures is a crucial to avoid a serious complications of stuck pipe situation. Stuck pipe avoidance measures can be categorized into preventative actions, that are taken to avoid the stuck incidents and mitigating measures applied when the situation has occurred (Albaiyat, 2012).

2.2.1 Stuck pipe prevention measures

It is of utmost necessity to put in place all measures to prevent the incidents, including proactive planning, diagnosis of stuck pipe situations, early sign detection and prediction approaches. Preventative measures are typically based on best practices and lessons learned gained from extensive drilling experience and managing stuck situations, they can be based on general practices or some location/field specific actions. Preventative actions usually encompass activities aimed to avoid stuck pipe hazard through established actions. Below is a review of several effective strategies commonly used in the industry (Dupriest et al., 2011; Mitchell, 2001; Heitmann & Burgos, 2015; Al Dushaishi et al., 2021):

BHA design: Using a stabilized BHA helps to minimize the contract area of drill pipe and the wellbore by incorporating stabilizers in the BHA. Replacing drill collars by Heavy Weight Drill Pipes (HWDP) in the BHA to enhance weight transfer to the bit and reduce the contact area between drill collar and wellbore. Placing additional accelerator alongside the jar in the BHA to amplify the impact force of the jar.

Rig operations: Reduction of pipe stationary time, Working the pipe before going for connection, Execution of wiper trips and reaming of tight intervals, and pumping specialized pills in the zones to erode mud cake (Drill and seal procedure) and pills to protect potential packing off zones (e.g., unconsolidated formations).

Mud properties: Maintaining required rheological mud properties, managing the solid content of the mud by eliminating high gravity solids to prevent the formation of thick mud cake, Regulating the ROP to prevent excessive strain on the mud circulation system, The measurement of torque and drag serves as an essential prevention method against mechanical sticking. Preserving the pressure overbalance at the lowest feasible value to avoid differential sticking

2.2.2 Stuck pipe remedy actions

Remedy measures greatly vary depending on the type of stuck pipe problem. In this context, detecting the early signs and mechanism that lead to the stuck pipe substantially increases the likelihood of releasing the drill string (Alshaikh et al., 2018). Over the years, the industry has evolved numerous procedure and practices derived from lessons learnt based on trial-and-error practices to address stuck pipe situations (Abbas et al., 2019). Mitigation measures can be classified into three primary categories depending on the cause of stuck pipe: pack-off and bridging, wellbore geometry induced stuck and differential sticking remedies. Nevertheless, the general strategy to release from any stuck situation is often similar in many cases, typically requiring interventions to modify the downhole condition carefully, and if unsuccessful apply substantial force to release it (Abbas et al., 2019). The first approach comprises pumping specially designed fluid containing lubricants, surfactants or fresh water down the hole in the annulus between the drill string and wellbore to alter the hydrostatic pressure, or to dissolve the cuttings that may cause the pipe to get stuck (Heitmann & Burgos, 2015). Different formations necessitate various conditioning mud types. For instance, salt formations require fresh water to dissolve it, whereas carbonate formations may require acid mud for dissolution of sloughed formations (Hilfiger et al., 2017).

The second way involves a force to be applied to the drill string when it becomes stuck. The force may be directed upward for pulling up or downward for slacking off, or rotational force by applying a torque. When the force exerted by the top drive is insufficient to release the stuck pipe, depending on its' type, depth, trajectory and formation, substantial impact force should be applied using specialized device known as jar. These devices normally integrated into the BHA and positioned within the drill string above the primary BHA components. To enhance jar efficiency, it is essential to note that jars are predominantly ineffective when positioned below the stuck point (Gonzalez et al., 2007).

Pack off and bridging remedy measures: Pack off and bridging induced mechanical sticking typically necessitates the application of impact force to release. It is crucial to ascertain that the feasibility of circulation is possible or attempt to regain the circulation by exerting downward force or by rotating the string. For the wells with high inclination the primary cause of packing off and bridging is accumulation of cutting beds, requiring the use of high-viscous pills to elevate the cuttings and transfer them to the surface. The jarring procedure is typically executed to free a string from mechanical stuck situations, involving multiple jar firing events that may extend from several minutes to several days. Normally, jarring is performed in the reverse direction of the pipe's last movement prior to becoming stuck (Mitchell, 2001; Alshaikh et al., 2018).

Wellbore geometry remedy measures: Stuck pipe incidents induced by wellbore geometry require force to be applied at the stuck point. Consequently, jar is fired in the opposite direction to the movement of the string prior to becoming immobile. A pumping procedure involving specially conditioned fluid becomes an option when jarring fails to solve drill string-stuck conditions at the wellbore area. Typically, it is a pill containing lubricating additives that are pumped into the borehole. Acids should be injected into shale or carbonate formations to dissolve the sloughed part that holds the pipe. Freshwater is used in salt deposits (Mitchell, 2001; Alshaikh et al., 2018).

Differential sticking remedy measures: The physical phenomenon underlying the differential sticking necessitates distinct activities to facilitate its release from a stuck situation. As previously noted, the difference between hydrostatic pressure in the drill column and formation pressure creates a pushing force which makes drill string stick to the wellbore wall (Reid et al., 2000). To liberate the string from its entrapment, it is essential to surpass the shear stress of the filter cake and the contact force between the drill string and the wellbore. Differential sticking usually occurs when the pipe remains stationary for extended duration (e.g., during connections when circulation ceases), whereas mechanical stuck pipe predominantly occurs during the movements of pipe in upward or downward directions (Dupriest et al., 2011). Considering all aforementioned factors, the actions pertaining to freeing the pipe aim to diminish the shear strength of the mud cake and decrease the force in the contact area between the pipe and the wellbore. Consequently, remedial measures commence by increasing the flow rate to the maximum allowable level to reduce the filter cake. The next phase involves slumping the string down or jar down while maintaining the half of makeup torque of the drill pipe connections. If unsuccessful, the subsequent remedial action should involve reducing the mud weight to alleviate the overbalance (Bowes et al., 1997). Nonetheless, this step should be executed with all the necessary precautions, as reducing the mud weight poses considerable risks when managing high pressure formations than can lead to fluid influx and kick. Hence, this measure is not used from the start. In this regard, if reducing the mud weight ineffective or deemed hazardous for operations, various combinations of spotting fluids have demonstrated their efficacy in releasing the pipe from differential sticking. The spotting fluids may possess various compositions, including lubricants, surfactants, mud cake dehydration additives, in certain instances, acids and solvents may also be incorporated. Their actions aim to lubricate the pipe and dry the filter cake to facilitate its cracking and release (Hilfiger et al., 2017). The mechanisms, root causes and prevention and mitigations measures for stuck pipe incidents are outlined in Table 2-1:

Table 2-1: Summary of stuck pipe mechanisms, causes, prevention and mitigation measures.

Type of stuck pipe	***Mechanical stuck pipe***		***Differential sticking***
Mechanism	Pack off and bridging	Wellbore geometry	Differential pressure
Cause	• Cutting's accumulation • Hole cleaning issues • Unconsolidated formations • Reactive formations	• Key seats • High doglegs • Mobile formations • Under gauge holes • Fractured formations	• High pressure overbalance in permeable formations • Increased stationary time in open hole • Formed thick mud cake
Prevention measures	• Optimized drill string design – stabilized BHA, HWDP • Maintain good hole cleaning		
	• Monitoring torque and drag readings • Controlling ROP to maintain mud cleaning process		• Reduction of pipe stationary time • Decrease the contact surface of drill pipe and formation • Use lower mud weight
	• Maintaining rheological mud properties • Regular hole cleaning • Sufficient flow rates to sustain optimal annular velocity	• Avoiding high doglegs • Reaming tight intervals, key seats, high dogleg zones • Drilling with higher mud weight salt intervals	
Remedy actions	• Applying impact force (jarring) • Use vibration tools • Rotating the drill string		• Increasing flow rate • Slumping or jarring down • Reduction of overbalance pressure • Pumping spotting fluids
	• Establish circulation • Pumping high viscous pills	• Pumping lubricants, fresh water (salts) or acids (carbonates)	

2.3 Drilling through evaporite formations

Oil and gas fields situated in the Pre-Caspian basin are characterized by their structural complexity, comprising substantial mobile Kungurian salts, represented by approximately 1200 salt dome structures, carbonate formations with thicknesses of 4000-4500 meters, and clastic rocks at the margins of the structure with thicknesses of 1000-2000 meters (Akhmetzhanov et al., 2020). Considerable challenges are associated with developing the field formed in the pre-salt hydrocarbon formations, as these formations developed beneath substantial salt domes that effectively served as trapping mechanisms for a large quantity of hydrocarbons. The formations exhibit a layered structure characterized by notable vertical and horizontal heterogeneity of the formation, low porosity and permeability in carbonates, the presence of anhydrite lenses, and a deep depositional depth of approximately 5000 meters. Drilling the wells in such a complicated environment brings significant challenges and necessitates specialized technologies, costly materials to ensure safe and efficient operations. This approach is essential for constructing robust wells capable of producing hydrocarbons over extended period of time. Despite the challenges introduced by presence of evaporite formations, they continue to serve as effective traps for massive hydrocarbon depositions around the world with successful commercial realization in Gulf of Mexico, North Sea, Brazil, West Africa basins (Omojuwa et al., 2011). The Pre-Caspian basin contains extensive hydrocarbon reserves that exist in pre-salt rock formations. This study investigates the field with deposits accumulated in pre-salt depositions at depths of 4000 to 5000 meters.

Evaporite formations typically presented in three main forms within oil and gas reservoirs (Omojuwa et al., 2011): Salt rock, known as halite, exists in the subsurface in two main forms: as either thin sedimentary layers or larger salt domes. Salt rock or halite, is typically soft and can occur as thin layers or substantial formations such as salt domes. Anhydrite represents the hardest type of salt formation, characterized by lower solubility and a more compact structural forms within the formation. Gypsum represents the hydrated state of anhydrite, which appears inside halite and dolomite subsurface structures.

The main challenges of drilling operations in salt formations, as shown in Figure 2-4, appear because of the unpredictable behavior of salts at different downhole environments. Under conditions of stress and elevated temperatures, salt has ability to plastically flow because of its creepy nature, leading to considerable wellbore stability issues, including stuck pipe, casing collapse, tightening the hole or lost circulation (Carcione et al., 2006).

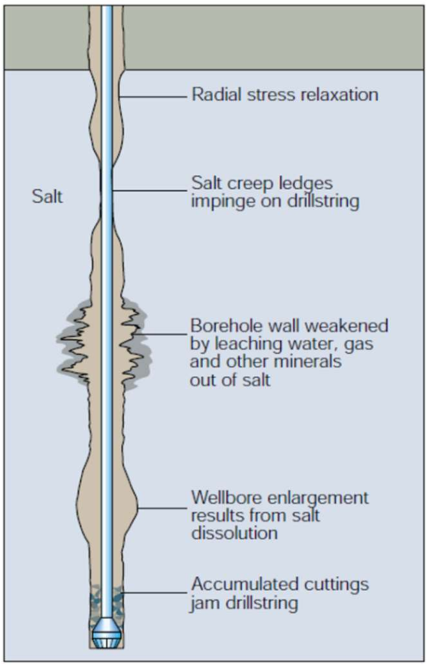


Figure 2-4 - Illustration of wellbore stability issues related to drilling salt formations (Campbell et al., 2021).

The creeping ability of salts is the primary cause of stuck pipe incidents while drilling. The ability of salt to plastically flow is influenced by two key factors: the composition of the salts and the temperature. Halite and anhydrite exhibit lower creeping properties, whereas carnallite demonstrates a high creeping effect. In addition to composition, temperature significantly influences the ability of salts to creep; at elevated temperatures, salts exhibit a greater tendency to flow. Furthermore, salts have higher thermal conductivity. This research paper focuses on the field formed during the pre-salt period, characterized by a deep depositional depth exceeding 4000 meters TVD and exposed to substantial downhole temperatures of more than 100 degrees Celsius. Which increases the salts mobility and the risks associated with potential stuck pipe.

As a result, effective planning is essential for drilling these formations. It is a standard practice to maintain a high mud weight when drilling through salt formations to prevent the salt formations from flowing into the wellbore (Dusseault et al., 2004). Figure 2-5 demonstrates the effect of mud weight on the creeping rate of various salt types. The graphs for halite and carnallite indicate that an increase in mud weight results in a reduction of hole closure attributed to the salt flowing effect (Campbell et al., 2021). This preventative measure also mitigates other substantial drilling challenges associated with casing collapse and cementing issues.

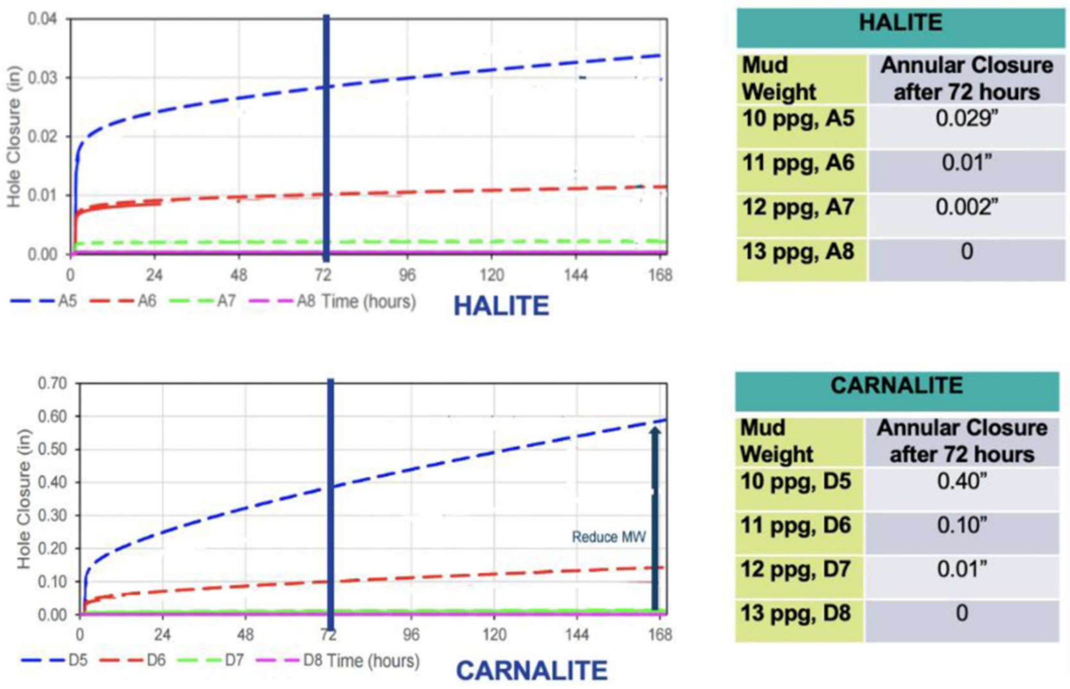


Figure 2-5 – Graphical illustration of creep rates of halite and carnalite salt formations under different mud weight conditions obtained from experiments (Campbell et al., 2021).

Moreover, there are significant risks associated when exiting or entering the evaporite formation. The formations adjacent to salt layers, above or below, are referred as a rubble zone (Amer et al., 2016). The inherent properties of salt complicate the prediction of formation pore pressure and fracture gradients in adjacent zones, resulting in considerable risks associated with blindly entering potentially high-pressure zones, which may lead to kicks, or low-pressure zones, which can cause substantial mud losses (Wang & Samuel, 2016). The lower limit of the mud weight window must account for the creeping behavior of salt to prevent it from flowing into the borehole, whereas the upper limit should not exceed the fracture gradient (Amer et al., 2016). The factors mentioned above reduce available safe mud weight window. Hence, maintaining the appropriate mud properties is challenging task when drilling through salt formation.

Another issue pertains to the salts' dissolution properties in drilling mud and cement, leading to hazards such as wellbore enlargements and contaminations of mud and cement. The industry predominantly uses salt-saturated water-based mud or oil-based mud to withstand the salt dissolution and maintain the stability of drilling fluids, particularly when drilling through extensive salt structures prolonged to thousands of meters (Mats Håpnes, 2014). To summarize, the oil and gas industry faces many difficulties when drilling through evaporite deposits. The prevention measures alongside planning activities do not eliminate the high risk of encountering stuck pipe incidents because of number of uncertainties of salt formation drilling operations. Hence, the accurate stuck pipe prediction approaches have the greatest interest among the industry, when dealing with such a challenging environment.

2.4 Stuck pipe prediction and detection approaches

The recent statistics has shown that the overall stack pipe incidents have significantly decreased over the years. Some researchers relate it to the active prevention measures used these days in the industry. A huge number of available best practices and lessons learnt, together with drilling personnel trainings bring notable changes in overall stuck pipe prevention measures. Some studies highlight that these improvement in the stuck pipe prevention also because of the active usage of automation in stuck pipe recognition which based on data-analysis using statistical approaches to define the stuck hazard indications based on drilling data (Alshaikh et al., 2019). Nevertheless, stuck pipe remains as one of the challenging sides of drilling industry, due the nature of the incidents, its complexity and uncertainties that follow the process. In this regard, stuck pipe prediction and detection approaches draw significant attention over the years.

In this chapter, the evolution of different prediction and detection methods are discussed in details highlighting the key industry approaches used these days. From the extensive literature review, prediction and detection models can be categorized into the three main broad categories, such as multivariate statistical analysis, analytical or physics-based approach, and ML-based intelligent models. Some researchers define two main categories, where all data-based models including statistical analysis and ML models combined into data-driven methods and second category is knowledge-based techniques (Zhang et al., 2019; Zhu et al., 2022; Kaneko et al., 2023). In addition, some hybrid approaches have been developed that incorporates the analytical methods together with data-driven models (Montes et al., 2023; Malki et al., 2024). These methodologies will be discussed in more details further in this chapter.

2.4.1 Multivariate Statistical Analysis

One of the earliest studies in using statistical analysis in stuck pipe recognition was performed by Lowe in 1983. Where historical data from 113 stuck cases in open hole from Gulf of Mexico was used to find the correlation between predictable variables, such as mud weight, inclination, and open hole in the form of stickiness factor, and the outcome variable of successful release from stuck, using least-squares regression curve.

The first studies towards the defining the stuck pipe from the data analysis while drilling has performed by Hempkins et al. in 1985. Their study incorporates the usage of multivariate discriminant analysis (MDA) to categories non-stuck, mechanical stuck and differential sticking conditions based on drilling data points. The research conducted its analysis using drilling data from 131 stuck pipe occurrences within Gulf of Mexico operations in the period from 1981 to 1984 by evaluating 20 relevant drilling parameters. The results indicated an accuracy of 81 to 87% in the classification of data among three groups.

The further study continued by Biegler and Kuhn in 1994, which enhanced the previous model by ability to identify the mechanism of pipe stuck incidents and define the stuck risk. The research included a dataset from 127 wells drilled in the Gulf of Mexico and used eight main drilling parameters. The classification conducted using MDA used Bayesian statistics to develop isoprobability curves to identify the risk level for future planned wells. A limitation noted in the paper is that the model is designed for water-based mud.

Application of multivariate statistical analysis continued in the same year 1994 by Howard and Glover with construction of two data sets from wells drilled in Gulf of Mexico and the North Sea. In their investigation, discriminant analysis exhibited noteworthy accuracy of outcome data with 75% when classifying into three categories of stuck condition: mechanical stuck, differential stuck and non-stuck; and 80% accuracy when classifying into two categories of stuck and non-stuck. They noted that accuracy of the model is significantly contingent upon the data set size and the input parameters. Developed model has the potential for real time application. Additionally, the application of MDA in predicting the stuck pipe freeing index was delineated.

The research performed by Shoraka et al. in 2011 demonstrated that the combination of multivariate statistical approaches, such as regression and discriminant analysis provide higher accuracy for stuck pipe prediction. The research used data from 40 wells in the southern oil reserves of Iran. A linear equation was formulated to predict differential pressure. Similar to previous studies, two discriminant functions are used to categorize the data into three groups: non-stuck, differential stuck, and mechanical stuck. The results indicated the accuracy of 84.5% in the classification of stuck pipe incidents. The primary limitation of the approach is its specificity for these particular fields that restricts the model to be directly used on the other locations.

Another research was related to studies using statistical analysis for pattern recognition to trigger warnings if the data is outside of the defined thresholds conducted by Weakley in 1990 to recognize the environment which has high risk of mechanical and differential sticking. The researchers have analyzed information from 600 different wells. Another study was done by Gulsrud et al. in 2009, where they used statistical analysis to assess the skew of bottom hole pressure (BHP) and stand pipe pressure (SPP), as well as the normalized standard deviation of torque (TRQ), resulting in a warning for inadequate hole cleaning or potential stuck pipe incidents.

2.4.2 Physics-based approach

The physics-based or analytical method is based on various physical phenomena that underlie the processes occurring while drilling. Theoretical solutions are compared with real-time data collected while drilling to ascertain the stuck pipe condition in drilling operations (Zhang et al., 2019). The industry predominantly uses Torque and Drag analysis (T&D), Drilling Hydraulics and Drillstring Mechanics models for well planning and real-time drilling operation diagnostics. The fundamental principle of physics-based approach often adheres the steps below (Cayeux et al., 2012): Continuous wellbore diagnostics achieved by the monitoring of real-time surface data. Comparison of actual acquired throughout the drilling operation with the physical models. Physical models' calibration to match the actual downhole conditions. Identify anomalous readings of actual parameters in comparison to the model. Inform relevant rig staff to initiate subsequent to prevent substantial drilling issues.

Typically, physics-based torque and drag analysis, along with drilling hydraulics are used for well planning and addressing certain engineering challenges while drilling. Nevertheless, they are not designed to be directly used in the real-time evaluation of the borehole conditions and prediction of drilling issues. The adoption of the models for real-time application is essential for effective monitoring of the well conditions (Cayeux et al., 2012). Hence, computing and monitoring friction factor is known as one of the key approaches for stuck pipe detection (Belaskie et al., 1994).

2.4.3 Torque and Drag analysis and Friction factor calculations

Stuck pipe refers to an immobile condition of the drill string down the hole, either partially or completely preventing movement in any direction. A variety of indicators are monitored by the driller to detect the sticking conditions during drilling operations. If the drag force, resulting from friction between the drill string and wellbore exceeds the anticipated force during pulling out of the hole, this condition indicates an overpull (Mitchell, 2001). The opposite force encountered while running in hole is referred to as slack off. If an overpull and slack off weights indicate zero value, then the drill string has no restrictions to move. Conversely, the changes in these values indicates that the hole getting tight, which may lead to a stuck condition. The process is done by monitoring the hook load (HL) readings at the rig site. During reaming operations, when the string is rotating, increase in torque readings at the stable RPM (rotation per minute) defines the presence of tight spots or possibility for sticking (Mohammed A., 2023).

Analysis of the drill string loads and their continuous monitoring is the critical step in designing and drilling the well. Torque and drag analysis is one of the common approaches for these purposes used in the drilling industry for decades. It was introduced for the first time by

Johancsik et al. back in 1984 and adopted into differential form by Sheppard et al. in 1987 (Mitchell & Samuel, 2009). As the name suggests, this analysis is used to calculate expected torque and drag impacting the drill string during various phases of drilling activity, including tripping, drilling or reaming. Drag force refers to the axial force acting between the drill string and the wellbore, while torque denotes the rotational forces occurring between the drill string and the wellbore. This analysis is performed through analytical calculations or numerical simulators, which are widely used these days (Aadnoy & Djurhuus, 2008). Real-time monitoring of torque and drag is commonly used to diagnose the pipe sticking issues (Lesage et al., 1988).

The model typically incorporates several input parameters, including BHA configuration and design, well trajectory, drilling mud characteristics and friction calculations between drill string and wellbore or casing (Elgibaly et al., 2017). The friction factor one of the main parameters in defining stuck condition of the drill string (Liu et al., 2024). Calculation of friction factor can be characterized by significant uncertainty, as it is not a measured parameter, but rather a combination of factors that affects the friction between the drill string and well bore, such as well trajectory, mud properties, particularly lubricity, pipe stiffness, drill string design, formation geology and accumulation of cuttings bed (Samuel, 2010). Two types of friction factor can be defined (Lesage et al., 1988): Friction factor between drill pipe and wellbore because of rotation (FRIC). Friction factor between drill pipe and wellbore while sliding (DRAG)

They are known as a Coulomb friction factor and defined as a ratio of friction force that acting on the pipe element to normal forces. The forces typically exerted on the pipe element include tension at both ends and buoyant weight of the element itself (Falconer et al., 1989). These forces generate a so-called side force at the contact area of drill pipe and wellbore, as depicted in Figure 2-6. The friction factor can thereafter be computed using the equations provided below (Belaskie et al., 1994):

FRIC (rotary friction factor) = Torque loss / (Side force x Element radius)   Equation 2-1

DRAG (axial friction factor) = Weight loss / Side force   Equation 2-2

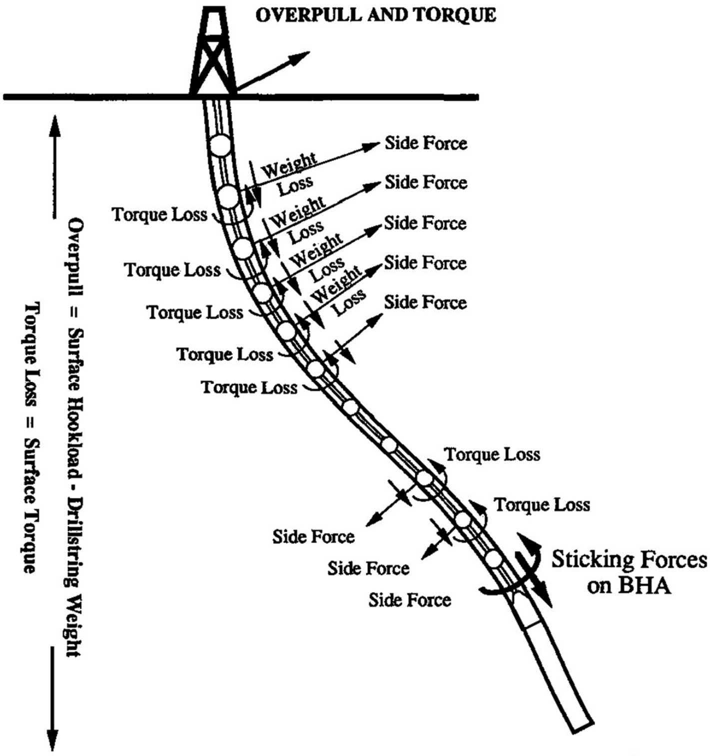


Figure 2-6 - Illustration of forces and losses that act on the drill string while off bottom (Belaskie et al., 1994)

Calculations of the friction factor obtained from torque and drug analysis are extensively used to identify the abnormalities in drilling processes, indicating potential stuck pipe issues. However, a number of limitations underlying the usage of friction factor-based models provides significant drawbacks. Physics-based model necessitates a substantial quantity of data to build a reliable model, typically encompassing but not limited to formation characteristics, mud properties, well trajectory and BHA design. In some cases, the absence of this data introduces considerable uncertainties that noticeably reduce the accuracy and reliability of the model (Kaneko et al., 2024). The proposed real-time friction calibration bases its approach on Cayeux and Daireaux's 2009 friction test strategy to reduce model uncertainty and improve its accuracy by performing regular friction factor calibrations and adjusting the model in real-time. However, the friction tests must be conducted regularly at specified intervals; this test is time-consuming and brings significant risks for pipe sticking, as it requires consecutive actions to move the string without circulation or rotation. In this regard, several researches have been conducted to address the limitations of purely physics-based models. As the result, numerous researches in recent times demonstrated increased interest in integrating both physical models along with data-driven approaches.

2.4.4 Hybrid approaches

The combination of models is aimed to reduce the uncertainty inherent in purely physics-based models, at the same time to enhance the data-driven models (Belaskie et al., 1994; Guzman et al., 2012; Salminen et al., 2016; Zhang et al., 2019; Zhu et al., 2022; Montes et al., 2023; Kaneko et al., 2024; Malki et al., 2024). In this approach, different methods incorporate analytical models with statistical tools and intelligent methods.

One of the early studies that used a combination of methods used physics-based models to monitor real-time data and its deviation from the modeled values (Belaskie et al., 1994; Guzman et al., 2012). Belaskie et al., proposed continuous real-time monitoring of friction factor to automatically identify anomies from modeled values and alert rig personnel. In the study of Guzman et al., the approach was enhanced by integrating the formation pressure monitoring, real-time Equivalent Circulating Density (ECD) tracking and Wellbore Stability and Mechanical Earth Model (MEM). This methodology introduced a so-called automated real-time modeling and analysis method (Ferreira et al., 2015; Salminen et al., 2016). Generally, this methodology predicts stuck pipe risks by detecting the difference between real-time drilling data with the results of analytical models of hydraulics and torque and drag analysis, which is achieved with the help of historical trends and domain expertise.

The most recent studies using a data-driven statistical method were performed by Zhang et. el., 2019. Where they used the hybrid model in conjunction with physics-based methods, they also used data-driven statistical tool known as the Ensemble Kalman Filter (EnKF) which compares model predictions with observational data. Authors claimed that the use of hybrid methodology yields more precise results, that can help to overcome the limitations of each method when used separately. The authors used Transient Solid Transport (TST), Torque and Drag (T&D) and Hydraulics and Mud Properties (HMP) models for physics-based analysis. The authors suggest that this method aids in differentiating the causes of stuck pipe, meaning that it can separate hole cleaning related issues from other formation or mechanical stuck pipe hazards.

Another hybrid approach was used integrating a physical model of torque and drag to derive the friction factor, which served as a distinguishing consideration for identifying stuck pipe condition. LSTM-based RNNs along with fuzzy mathematics operated together to create a wide-range stuck pipe risk analysis (S. Zhu et al., 2022). An accuracy level of 0.98 was achieved through results and the automated system kept false alarm occurrences less than 1%. Authors conclude, that physical models usually lack precision and do not facilitate early detection of stuck pipe, whereas the pure LSTM model exhibits a significant false alarm rate. Hence, the combination of approaches aids in overcoming these limitations.

Torque and Drag model remain as a primary method for physics-based models; however, in some researches the authors also used other physical models to enhance the results, such as Cuttings Transport Model (Zhang et al., 2019). Moreover, the analysis combined ECD with K-Nearest Neighbors (KNN), Logistic Regression (LR), Multi-Layer Perceptron (MLP), RNN and Gaussian Naive Bayes (GNBC) approaches for predicting early stuck pipe signals in high risk well regions with 4-hour incident prediction capabilities (Malki et al., 2024). A research team used Geometric Sticking Analysis, specifically Tortuosity & Rugosity Index, combined with Unsupervised Fuzzy Clustering in Utah FORGE Geothermal wells to identify the primary sticking reasons that occur when drilling shifts from sliding to rotating modes, primarily through high-tortuosity areas. The result demonstrated an accuracy of 87%, with a false alarm rate of 11%, and ability to detect the early warnings 1-3 hours before the occurrence (Montes et al., 2023).

2.4.5 Machine Learning Approaches

With development of high-performance computers, ML algorithms were adopted for petroleum engineering downstream sector. Further development of data-driven approaches showed transition from statistical models towards ML-based intelligent models.

One of the earliest studies in using ANN for stuck pipe prediction was done by Siruvuri et al. in 2006. The authors used a neural network approach to both predict and mitigate the occurrences of differential stuck pipe incidents. The authors propose a proactive solution that adopts a feed-forward neural network with backpropagation trained on 200 datasets, including 120 stuck cases and 50 non-stuck cases, that involved differential pressure parameter, mud properties, such as viscosity, fluid loss and drilling operational conditions. Distinct models were created for water-based and oil-based mud systems. The error metrics evaluated the model with moderate precision, MSE ranging from 0.01 to 0.9, this performance relied significantly on the quality of input data and selected parameters. The authors emphasize the capability of the model for real-time monitoring using drilling data to proactively mitigate risks of stuck pipe.

Another early application of ML for stuck pipe prediction purposes was performed by Miri et al. in 2007 to predict differential pipe sticking during drilling operations in Iranian offshore oil fields. The research used a database from 64 wells in the Persian Gulf, including data from both stuck and non-stuck pipe incidents. Prediction model was created based on two types of neural networks – Multi-Layer Perceptron (MLP) and Radial Basis Function (RBF).

The backpropagation algorithm used to train the model through cross-validation and verification on testing datasets. In addition, the authors performed sensitivity study to access the influence of each input parameter on the output using ANN.

In the research conducted by Murillo et al. in 2009 in conjunction with ANN adaptive fuzzy logic was used to improve the accuracy of the results. Total of 185 datasets with number of input parameters were used for the study. In their research discriminant analysis was used to classify the data into three categories of stuck pipe, as mechanical stuck, differential sticking and non-stuck, while ANN was used to detect the non-linear relationship in the input data. At the same time adaptive fuzzy logic was used to enhance the discriminant analysis and ANN approaches. According to the authors, the combination of approaches yielded an accuracy of 98% in the output results.

A number of studies have been conducted in the area of pure ANN based models with the use of feed-forward neural networks (Shadizadeh et al., 2010; Nezhad et al., 2012; Jahanbakhshi & Keshavarzi, 2016; Q. Zhu et al., 2019). These researches have used the models designed and adopted mainly for specific fields where data was obtained from, demonstrating the high accuracy of results, however, with limited generalization of the models.

2.5 Intelligent Prediction Models

Many researchers demonstrate growing interest in using different ML-based intelligent prediction models in their studies. From the way how ML-based models learn from data they can be categorized into three main groups: supervised learning, unsupervised learning and reinforcement learning. Where supervised learning covers a brought group of algorithms which uses labeled input data to make a prediction of output data. Supervised learning is dealing with both classification problems, where output is in discrete form of categories, and regression cases, where output represents a continuous form of variables (Bishop, 2006). While in unsupervised learning, input data is used without any labels and ML-based models learn the patterns without any supervision.

The focus of this research work is on stuck pipe prediction which represents a classification problem, as desired output of the modeling is expected to be in a discrete form of categories of stuck pipe or non-stuck. Hence, in this literature review supervised ML classification algorithms will be discussed in more details.

2.5.1 Supervised machine learning algorithms

Supervised ML technique which demonstrated high accuracy and reliable outputs is known as Support Vector Machine (SVM). This algorithm can deal with both classification and regression problems, and it is suitable for high-dimensional datasets especially in binary classification purposes. SVM algorithm was introduced for the first time by Vapnik in 1995. SVM uses the hyperplane to define the best decision boundary for the better separation into classes, which makes this algorithm one of the powerful tools for classification problem (Rostami & Manshad, 2014). One of the first use of Support Vector Machine (SVM) algorithms performed by Jahanbakhshi et al. in 2012. This research evaluates performance of SVM and ANN algorithms in determining differential pipe sticking rates in Iranian offshore oil fields. The authors validated their model by training it on 70% of the data and testing it on 30% through analysis of drilling parameters in 214 samples from 63 different wells. The SVM technique equipped with a Gaussian kernel is Radial Basis Function (RBF) Kernels known as distance-based similarity functions, delivered more accurate results than the ANN model providing 92.19% accuracy combined with 93.75% sensitivity and 90.63% specificity while ANN produced 82.81% as its maximum accuracy.

Kernel function performs similarity calculations for data points rather than computing all features. The training process in high-dimensional spaces becomes more efficient through this method without running extensive calculations. Kernel matrix (also known as Gram matrix) contains all data point pairwise similarities as the sole requirement for operation (Albaiyat, 2012). SVM provides three main benefits to its users: higher generalization capability alongside minimizing overfitting, and its ability to work with limited number of data sets.

The other very first use of SVM in predicting the stuck pipe was done by Chamkalani et al. in 2013. A new approach to drilling operational stuck pipe prediction develops a hybrid Least Squares Support Vector Machine (LSSVM) model which uses Coupled Simulated Annealing (CSA) optimization method. An evaluation of the model on the Iranian oil field data revealed that it reached 94% success in detecting the difference between stuck and non-stuck states. Overall, other comparison studies on the use of SVM with ANN indicated a higher accuracy of SVM models compared to ANN models (Jahanbakhshi et al., 2012; Albaiyat, 2012; Rostami & Manshad, 2014; Alshaikh et al., 2019; Abbas et al., 2019).

Use of ML algorithms brought further adoption and use of various models based on ML algorithms for stuck pipe detection and prediction. Especially supervised learning algorithms are widely used by many researches in their study of stuck pipe detection because of a number of advantages, including accurate predictions with quantifiable outcomes, transparency in decision-making processes, and sophisticated algorithms capable of capturing complex patterns in the data (Magana-Mora et al., 2019; Alshaikh et al., 2019; Ahmed et al., 2019; Elmousalami & Elaskary, 2020; Al Dushaishi et al., 2021; Mal et al., 2022; Katzmann et al., 2024).

One of these algorithms is known as Decision Tree (DT). This algorithm is part of supervised ML techniques, known as fast and robust tool for presence of outliers (Hastie et al., 2009). As the name suggests, the architecture represents a tree alike form, where the test features correspond to internal nodes, the tree's branches represent possible output of the model, and its leaf nodes represents results of classification (Saporetti et al., 2018). This algorithm is widely used for ability of internal parameter selection that makes the model robust for number of unnecessary parameters, its transparency and explainability (Antonini et al., 2024). DT together with Random Forest (RF) approaches demonstrated good efficiently and accuracy in the study performed by Magana-Mora et al., in 2019. The highest accuracy of the output of 82.75% was obtained with the use of RF algorithm compared to ANN and SVM. The model efficiency was additionally improved by using statistical tools for feature selection prior to training the models. Alshaikh et al. in their study additionally used statistical techniques to identify the critical drilling parameters for model enhancements. The SVM model achieved 98.62% accuracy based on a study that included ANN, SVM and DT algorithms. Nonetheless all models demonstrated outstanding results with accuracies surpassing 96%. Both studies demonstrate that selection of input parameters represents a crucial aspect that determines output results. Consequently, data pre-processing with defining the key parameters is a crucial stage in successful use of the models, as it greatly aids in analyzing the root causes of stuck pipe incident and determining appropriate preventive actions.

Some researchers have conducted studies which used different classification methodologies including Logistic Regression (LR) and K-Nearest Neighbors (KNN) over the past decade (Elmousalami & Elaskary, 2020; Malki et al., 2024). Logistic Regression (LR) algorithm known as one of the main data analytical tools used for defining the relationship between dependent variable and one or more independent variables. Usually, the output of this model is presented in discrete form as binary output of two or more values in the case of multiple classification tasks. In comparison with LR algorithm, which is known as one of the commonly used data analysis tools that provides continuous output, LR is used for classification problems, as it predicts probabilities and provides the results in discrete categories form (Hosmer & Lemeshow, 2000). Depending on the number of outputs it can be categorized into binary with two outputs, multinomial with several outcomes, or ordinal logistic regression with ordered classes.

K-nearest neighbors (KNN) is supervised ML algorithms that can be used for both classification and regression problems. This approach is considered as simple but powerful tool (Saporetti et al., 2018). Unlike other ML-based models, KNN does not learn from the data patterns, but rather memorizes the patterns to make a prediction by comparison of stored data with the new one. Hence, this algorithm also known as instance-based learning algorithm. Despite the simplicity of the method, it demonstrates quite high accuracy and predictive capabilities in more complex cases (James et al., 2021). However, notable drawback of storing the training data may create additional issues with increased computational time and requirements for storage of big amount of data. Therefore, this method is suitable for small datasets. At the same time, KNN, similarly to the LR, is sensitive to outliers which may lead to misclassification of data, meaning that preprocessing step is critical for this technique (Wu et al., 2008).

Moreover, an ensemble learning algorithms, which integrates multiple models, such as boosting algorithms, have shown powerful results in number of studies (Ahmed et al., 2019; Elmousalami & Elaskary, 2020; Brankovic et al., 2021; Malki et al., 2024). Ensemble learning algorithms combines several learners and their predictions to develop powerful model that overcomes the main disadvantages of multiple approaches to improve the performance and accuracy (Sewell, 2011). This algorithm allows to combine several ML algorithms such as DT, SVM and other boosting methods with ANN (Elmousalami & Elaskary, 2020). Depending on number of classifiers used in the algorithm there are different combinations of ensemble learning models can be found (Sewell, 2011). The most common approaches include bagging, boosting, stacking and voting (Elmousalami & Elaskary, 2020).

Boosting algorithms developed over last decades and highly used as intelligent methods, they are known as a powerful algorithm that combines several “weak” classifiers into high performance model. Originally, boosting approach was used for classification problems, however, later it was extended to regression tasks (Hastie et al., 2009).

2.5.2 Unsupervised and deep learning approaches

Advanced ML techniques and deep learning algorithms combined with unsupervised learning for stuck pipe prediction and early signs detection have been extensively researched by many scholars (Inoue et al., 2021; Mopuri et al., 2022; Tsuchihashi et al., 2021; Meor Hashim et al., 2021; Othman et al., 2022; Elahifar & Hosseini, 2022; Mal et al., 2022; Malki et al., 2024). Algorithms of unsupervised learning, in comparison to supervised learning, operate on unlabeled data to define the patterns among parameters, as this technique functions independently of pre-defined input and output labels. Some researchers contend that unsupervised learning has greater potential for identifying early signs of stuck pipe situations compared to supervised learning which learns the stuck pipe conditions rather than early signs. Nonetheless, unsupervised learning has certain limitations, such as lower accuracy of the output data because of variability of input drilling parameters (Kaneko et al., 2024).

In the research conducted by Tsuchihashi et al., 2021 a deep learning method 3D – Convolutional Neural Network (CNN) was used to detect the early signs of stuck pipe condition. In comparison with other studies which used time-series analysis, the authors utilized depth-domain method. The project used the data from 30 distinct fields encompassing 19 key drilling parameters. As a result, their model demonstrated the higher accuracy based of AUC score of 0.6527 of depth-domain 3D-CNN over time-series model with AUC score of 0.475. The authors believe that their methodology enhances the early detection of stuck pipe and reduces excessive false positive warnings.

Another advanced deep learning algorithms used for detection of early indicators of stuck pipe incidents known as Long Short-Term Memory (LSTM) Networks, which is a subset of Recurrent Neural Network (RNN) approaches. LSTM demonstrated very good performance in handling sequential and time-dependent data generated during drilling operations, and proved as an effective method for time-series forecasting (Mal et al., 2022; Othman et al., 2022). In the study of Mal et al., the highest accuracy of 98.99% was attained using LSTM in conjunction with RF classification approaches. While research performed by Othman et al. in 2022 showed enhanced predictive capabilities when integrated with ANN by identifying the risk of stuck pipe because of geometry issues 6 to 10 drilling stands ahead.

Multiple studies demonstrated that deep learning with unsupervised algorithms create better performance outcomes and new possibilities for early stuck pipe signs detection (Inoue et al., 2021; Mopuri et al., 2022). Inoue et. al., in 2021, used unsupervised deep learning model to detect the early indicators of stuck pipe prior to their occurrence. Authors propose a technique to assess standard drilling operations and identify data anomalies related to stuck events. The model deploys deep learning techniques through Autoencoders linked with Long Short-Term Memory (LSTM) that analyze drilling data patterns, alongside Probability Mixture Models for stuck condition prediction, and Gating Module to pick the optimal model for different drilling operations. The study used the data from 73 wells sourced from various agencies. The model was assessed using Receiver Operating Characteristic (ROC) curve and Area Under the Curve (AUC) metric. Their model demonstrated a superior accuracy of 0.79 AUC score over supervised learning model. In the study of Mopuri et al., 2021 the AUC score for unsupervised deep-learning model outperformed supervised approaches such as SVM and RF.

Another intelligent method that also used in stuck pipe prediction studies was application of Reinforcement Learning (RL) algorithm alongside with other ML methods (Alzahrani et al., 2022). RL algorithm differs from supervised learning, as it does not assign the labels to input and output data, instead it uses the trial-and-error method to improve the decision-making process. The study develops stuck pipe prediction approaches by treating it as a multi-class classification task which enhances the model's prediction accuracy and adaptability when drilling conditions change. The model classifies the data into six classes, according to ongoing operations (e.g., run in hole, drilling, pull out of hole). The authors indicate that, the model continuously self-enhancing with the use of through RL, attaining an accuracy of 98.5%.

In addition to prediction and detection approaches of stuck pipe, application of ML techniques for determining remedy measures have gained popularity among certain researchers (Abbas et al., 2019; Elmousalami & Elaskary, 2020; Al Dushaishi et al., 2021). The same methodology used for detection was applied in identifying correct mitigation measures assessed by ANN and other ML algorithms. SVM based models demonstrated the higher accuracy over ANN model in determining appropriate remedial procedures (Abbas et al., 2019). In the study of Al Dushaishi et al., 2021, recursive partition and DT analysis was used to predict stuck pipe incidents and propose mitigation strategies for each case, including working the pipe, applying chemicals (e.g., HCl acid, caustic soda), using a jar, or performing a sidetrack. The authors reported an 84% accuracy in predicting suitable remedial actions. A summary of the literature of the main ML approaches applied for predicting and detecting stuck pipe incidents along with defining the remedial measures is presented in Table 2-2.

Table 2-2 – A summary of the literature on main stuck pipe prediction and detection intelligent models.

Method	Model details	Dataset details	Input parameters	Results / Performance	Remarks	Authors
Multivariate Statistical Analysis						
MDA	- MDA - Bayesian statistics - Multivariate statistical regression	- Gulf of Mexico wells (1981–1984, 127–131 cases), - North Sea wells - 40 wells Iranian oil fields	- Main drilling and mud properties	- 75–87% accuracy - Classification into stuck/ non-stuck, mechanical stuck, differential sticking groups	- Model designed for water-based mud (Biegler and Kuhn, 1994) - Applicable for a specific field (Shoraka 2011)	Lowe1983, Hempkins 1985, Biegler and Kuhn 1994, Howard and Glover 1994, Shoraka 2011
Threshold-based Warnings	- Skew analysis (BHP, SPP), - Normalized standard deviation (TRQ)	- 600 wells (Weakley 1990), - 3 wells of Iranian fields (Gulsrad 2009)	- Standpipe pressure and surface torque (Gulsrad 2009)	- Pattern recognition in high-risk environment	- Requires predefined thresholds - Limited generalization	Weakley (1990), Jardine (1992), Gulsrad (2009)
Machine Learning (ML)						
ANN models	- Feed-forward Neural Networks, - Multilayer Perceptron (MLP) - Radial Basis Function (RBF)	- Persian Gulf wells (64–200 datasets), - Iranian oil fields	- Main drilling and mud properties	- 82–98% accuracy	- High dependency on data quality - Overfitting risks - Applicability for a specific field	Siruvuri (2006), Miri (2007), Murillo (2009), Shadizadeh (2010), Albaiyat (2012), Nezhad (2012), Jahanbakhshi and Keshavarzi (2016), Zhu (2019)
Supervised ML	- SVM (Gaussian kernel) - Random Forest (RF) - Decision Tree (DT)	- Gulf of Mexico wells - Iranian oil fields - Global datasets (30–214 samples)	- Main drilling and mud properties - Well trajectory data - BHA design	- RF accuracy 82.75% - SVM accuracy 92–98.62%, - Remedey measures accuracy up to 84% (DT)	- SVM outperforms ANN in generalization - Input data selection is critical	Jahanbakhshi (2012), Albaiyat (2012), Chamkalani (2013), Magana-Mora (2019),

	- Logistic Regression (LR) - K-Nearest Neighbors (KNN), - Boosting algorithms					Alshaikh (2019), Ahmed (2019), Elmousalami and Elaskary (2020), Al Dushaishi (2021), Brankovic (2021), Mal (2022)
Unsupervised / Deep Learning	- Autoencoder-LSTM - 3D-CNN - RL - Fuzzy clustering	- 73 global wells - Middle East oils fields	- Main drilling and mud properties - Well trajectory - BHA design	- Autoencoder-LSTM AUC 0.79 - 3D-CNN AUC 0.6527, 87% - Fuzzy clustering accuracy 87%	- Lower accuracy vs. supervised approaches - Intensive computations	Inoue (2021), Tsuchihashi (2021), Mopuri (2021), Othman (2022), Elahifar and Hosseini (2022), Mal (2022), Malki and Abughaban (2024)
Physics-based approach						
Torque and Drag (T&D)	- Friction factor analysis - Hook load monitoring	- Global wells	- Real-time surface and drilling parameters	- Detects tight spots via torque/drag deviations	- Requires frequent calibration - Uncertainties in friction factor calculations	Johancsik (1984), Sheppard (1987), Lesage (1988), Belaskie (1994), Cayeux and Daireaux (2009), Mitchell (2011), Mohammed A. (2023)
Hydraulics/Mechanics	- ECD analysis - Cuttings Transport Model (TST)	- Global wells	- Real-time surface and drilling parameters	- Early warnings for hole cleaning issues	- Limited to specific operations - Requires high volume of data	Cayeux (2012), Zhang (2019), Malki and Abughaban (2024)

Hybrid Approaches						
Physical and Statistical	- Friction factor analysis - ECD analysis - FPWD and MEM - Pattern recognition	- Gulf of Mexico (Belaskie, 1994) - Saudi Aramco wells (Guzman, 2012)	- Hookload, surface torque, downhole weight-on-bit and torque (Belaski, 1994)	- Early warnings for stuck pipe conditions - Reduced false alarms	- Uncertainties in physical models - Limited generalization	Belaskie (1994), Guzman (2012)
Physical and ML/Statistical	- EnKF (Ensemble Kalman Filter), - LSTM w/ fuzzy logic, - ECD analysis w/ KNN and LR	- Global wells - Geothermal wells datasets	- Main drilling and mud properties - Well trajectory - BHA design	- LSTM + fuzzy F1 score 0.98 - LSTM + fuzzy model – 4-hour early warnings - ECD + ML – 1–3h early warnings	- Combines model strengths - Reduced false alarms (e.g., 1% in Zhu 2022)	Zhang et al. (2019), Zhu (2022), Montes (2023), Kaneko (2024), Malki & Abughaban (2024)
Physics + Unsupervised ML	- Geometric sticking analysis (tortuosity/rugosity) w/ fuzzy clustering	- Utah FORGE geothermal wells	- Main drilling and mud properties - Well geometry and trajectory - BHA design	- 87% accuracy, - 11% false alarm rate	- Focuses on geometric root causes - Model is not limited to geothermal wells	Montes (2023)

3. METHODOLOGY

The input data and methodology used to conduct the research are presented here in this chapter. The main objective of this research work was to develop the most accurate and reliable stuck pipe prediction intelligent models based on ML algorithms such as Logistic Regression, K-nearest neighbors (KNN), Support Vector Machine (SVM), Decision Tree (DT) and CatBoost – gradient boosting algorithms. Information on data collection including surface drilling data acquisition and input parameters selection processes is discussed in part one followed by data pre-processing steps required before data modeling including data cleaning, detection of outliers, and missing data. The ML approach used to construct proposed intelligent stuck pipe detection models are presented in the third part. And finally, model performance evaluation metrices are presented by using data accuracy and error calculations.

3.1 Data Collection

As data-driven approach, Machine Learning algorithms learn from data to define the patterns and perform a reliable prediction. Hence, the quality, consistency and reliability of the data are of significant relevance when using ML algorithms. This research used historical data from previously drilled wells in one field to obtain a data set from 10 wells. Data collection was one of the main phases of the project, characterized by notable time-demanding and challenging nature, requiring the acquisition of information from multiple sources, which include different reports, time and depth logs, and various data files. Moreover, there is not much interest among many oil companies to collect and store unused drilling data (e.g., time-based data) because of substantial volume and time-intensive tasks. Consequently, various approaches of data extraction along with available information were used in this research to achieve the highest outcome from the project.

3.1.1 Data acquisition

Various operating companies use their own approach on data acquisition while drilling operations. Nevertheless, some unified approach in the industry exists, which usually incorporates the following data:

Mud logging data presented by: drilling parameters obtained from surface sensors installed at the rig. Geological information based on cuttings analysis using different techniques. Gas analysis and oil-bearing zone evaluation approaches and geological characterization. Drilling fluid data, including number of different physical and rheological properties, obtained by the standardized laboratory measurements and sensors installed at the rig. Downhole drilling data (MWD and LWD), obtained with the use of various downhole equipment as part of BHA. Geological and geophysical information that collected from LWD and wireline formation evaluation technologies. Reservoir and production data.

In this research work, required data was collected from number of different sources: Mud logging time logs, mud logging master logs, geological reports, drilling fluid reports (DFR), daily drilling reports (DDR), end of well reports (EOWR), and BHA report.

During analysis, it was decided to use time-based mud logging data, as some of the stuck cases occurred during making connection while the pipe was stationary, recorded at 5 seconds interval. Number of challenges associated with data collection include inconsistences in data sources and sampling rates; for instance, time log data is recorded every 5 seconds in contrast with information given in DDR, which is reported every 15 minutes. Another example is drilling fluid data, which is usually collected from laboratory measurements conducted every 6 hours during critical operations, such as drilling, and every 12 hours in the case of other activities.

3.1.2 *Field data analysis of stuck pipe condition*

Data collection was conducted from ten wells drilled on X field, the real field and well names, as well as the Company name will not be published for confidentiality reasons. The wells have vertical profile and targeted deep depositions with true vertical depth (TVD) of around 5,000 m. The wells were drilled in the challenging environment of Pre-Caspian basin targeting sub-salt formations with number of geological and engineering challenges. The challenges include drilling shallow unconsolidated formations, the zones with layered structure of carbonate rocks (e.g., limestone) and terrigenous formations (e.g., sandstone), characterized by notable vertical and horizontal heterogeneities, hard drilling intervals (HDI) presented by anhydrite formations, and drilling massive salt structures presented as salt domes. Wellbore stability was a main concern of drilling these wells, as number of stuck pipe cases were exhibited while drilling and notable mud losses intervals observed. The main driver for the project was the Company's concern regarding the future field development plan, that include additional new wells, and exploration of new solutions for stuck pipe prediction problem. The collected data consists of input parameters obtained from 10 wells, incorporating 61 data sets with a total of 610 data points, including 30 – non-stuck cases and 31 – stuck cases. Typical well profile is presented in the Figure 3-1 and include the following sections (Table 3-1):

Table 3-1 – Hole size and Casing data:

Bit Size (inch)	*Depth (m)*	*Casing type*	*Casing Size (inch)*	*Depth (m)*
26"	300	Surface casing	20"	300
16"	1250	1st Intermediate casing	13 3/8"	1250
12 ¼"	3220	2nd Intermediate casing	9 5/8"	3220
8 3/8"	5200	Production Casing	7"	5200

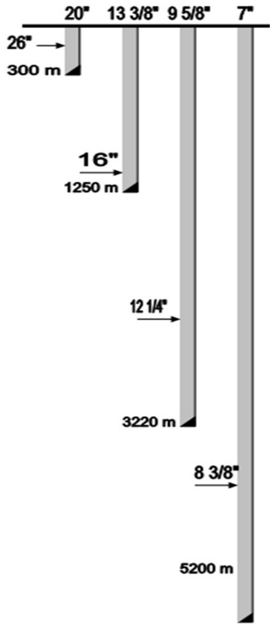


Figure 3-1 – Casing design of the X Field wells

While analyzing the data for stuck pipe occurrences, the following observations were made: Stuck pipe situations took place during drilling operations across various hole sections, including top sections and challenging salt intervals. Stuck pipe incidents have diverse nature and causes, induced by number of reasons including geology and formations, drilling activity, equipment related issues. Majority of stuck pipe incidents occurred during reaming and tripping activities, such as reaming up or down, pulling out of hole (POOH) or running in hole (RIH), with several cases recorded while drilling and when the pipe remained stationary for some period of time (e.g., connection). Along with stuck pipe incidents, all instances of pack off, tight spots and overpulls were evaluated, to define the early signs of stuck pipe occurrences and their prediction. The following table summarizes data classification based on drilling activity and type of stuck pipe condition (Table 3-2):

Table 3-2 – Input data classification based on activity and stuck type:

Activity	**Drilling**	**Reaming**	**Tripping In / Out**	**Connection (stationary)**	**Total number**
Stuck pipe	3	9	1	1	14
Tight spot	-	1	1	-	2
Overpull	-	1	4	-	5
Pack off	-	8	2	-	10
Total Stuck	3	19	8	1	**31**
Non-Stuck					**30**
Total					**61**

Further analysis has been performed to determine the correlation between stuck pipe incidents and the hole section drilled. As expected, it was noticed that the main stuck pipe instances occurred during drilling intermediate hole sections within salt intervals, followed by transition intervals with layered formations, with only few incidents recorded while drilling top sections of unconsolidated formations and production interval. Data classification based on drilled hole sections is summaries in the Table 3-3.

Table 3-3 – Input data classification based on hole section and stuck type:

Hole Size (inch)	**26”**	**16”**	**12 ¼”**	**8 3/8”**	**Total number**
Stuck pipe	1	3	8	2	15
Tight spot	-	-	2	-	2
Overpull	-	1	4	-	5
Pack off	-	9	1	-	10
Total	**1**	**13**	**15**	**2**	**31**

3.1.3 Input data selection

As it was highlighted in the previous chapters, selection of input data is of the utmost significance for accurate and early stuck pipe detection approaches (Siruvuri et al., 2006; Murillo et al., 2009; Albaiyat, 2012; Magana-Mora et al., 2019; Alshaikh et al., 2019). Required parameters should correlate with conditions that can depict the early signs of stuck pipe incidents, at the same time these specific parameters should be consistently measured across different rig activities and recorded at sufficient sampling rate (Brankovic et al., 2021)). In this regard, mud logging data serves as the primary source of input data required for analysis with the set of additional information that significantly influences stuck condition, such as drilling fluid information, well profile data, BHA design and lithology, which notably differs depending on hole section drilled. Total of 24 input parameters were used to train and test the models. All input parameters are presented in the Table 3-4 with more details.

Table 3-4 – Input parameters details:

Data category	**Source of data**	**Input parameter**	**Acronym**	**Unit**	**Description**
Drilling Parameters	Mud Logging data and Daily Drilling Reports	Bit depth	DPTS	m	Measured depth of the bit
		Rate of Penetration	ROP	min/m	Drilling rate of penetration
		Weight on Bit	WOBM	tons	The amount of surface weight applied to the bit
		Revolutions per minute	RPM	rpm	Rotating speed of the BHA
		Surface Torque	TORM	kgm	Surface rotary torque applied to the BHA
		Hook load	HKLA	tons	The load on the hook
		Stand Pipe Pressure	PSPM	atm	Pressure of the mud in the stand pipe
		Mud flow in	MFIN	l/min	Mud flow rate
		Mud weight	MW	kg/dm^3	Mud weight

Mud Properties	Drilling Fluid Report	Marsh Funnel Viscosity	MFV	sec/L	Funnel viscosity of the mud
		Plastic Viscosity	PV	cp	Plastic viscosity of the mud
		Yield Point	YP	cp	Mud's yield point
		Gel Strength	GEL10SEC	$lbf/100ft^2$	Shear stress of the mud measured after 10 seconds
		Gel Strength	GEL10MIN	$lbf/100ft^2$	Shear stress of the mud measured after 10 minutes
		Solid Content	SOLID	%	Total solid content
Lithology	Mud logging data	Salt content	SALT	%	Percentage of salt content
		Clay content	CLAY	%	Percentage of clay content
		Anhydrite content	ANHYDRITE	%	Percentage of anhydrite content
		Sandstone content	SANDSTONE	%	Percentage of sandstone content
		Limestone content	LIME	%	Percentage of limestone content
		Other lithologies	OTHER	%	Percentage of other lithologies
Other Information	BHA Report and EOWR	BHA length	BL	m	Total length of BHA components
		Hole Size	HS	in	Hole section size
		Activity	Name	n/a	Current activity name: Drilling, Reaming up / down, POOH, Connection

3.1.4 Stuck pipe condition description

Certain behavior of drilling parameters is expected when dealing with stuck pipe situation. During normal drilling activities without any anomalies measurements of the drilling parameters remain within a certain range. As illustrated in the time log sample (Figure 3-2), selected time interval represents a normal drilling activity.

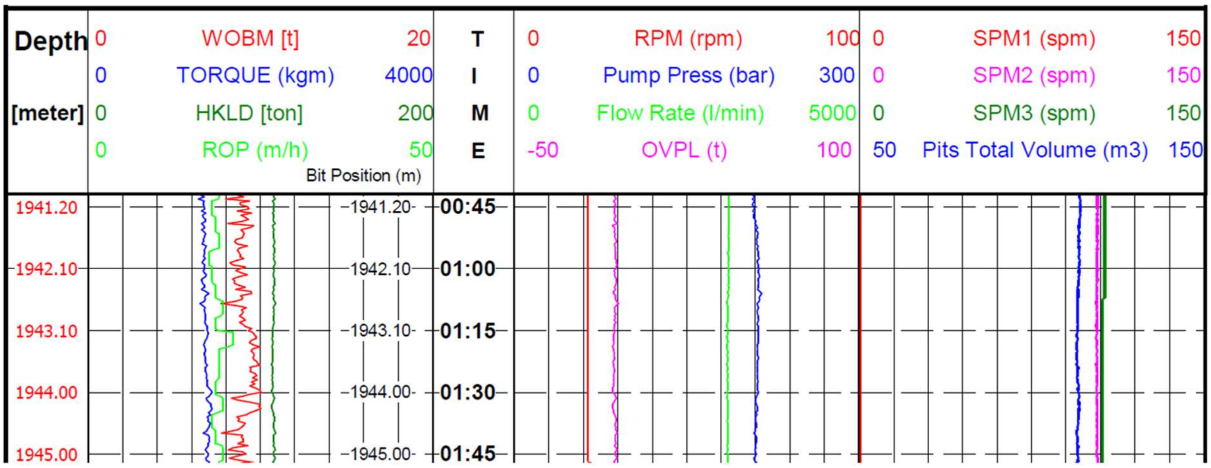


Figure 3-2 – Example of time log interval with regular drilling activity.

The depth measurements increase as drilling advances, no significant fluctuations of drilling parameters observed. While Figure 3-3 shows a time log segment when the pipe became stuck during drilling. In comparison to normal drilling activity, it is clearly seen that the curve patterns noticeably deviate from their regular conditions: The bit position is not changing or shows limited movement of string with some driller's attempt to pull up the string. Torque (TORQUE) values show erratic and spiky behavior with fluctuations of 0 values to significant torque increase, showing the string's inability to rotate. Similar behavior of rotating speed (RPM) to torque readings observed, the values drop to 0 with spiky readings. Pump pressure (Pump Press) and flow rate (Flow Rate) also demonstrate notable changes indicating some blockage in the mud flow. Hook load (HKLD) readings fluctuate markedly demonstrating inability to move the string up, with several unsuccessful attempts of driller to apply some load to release the pipe.

These are the common characteristics of a stuck condition; however, they are not uniform for all stuck cases. Behavior and patterns of parameters may vary depending on the stuck condition, its mechanism, geology and trajectory, current activity and many other factors. Notwithstanding, similar behavior of two to three drilling parameters is common in the most of stuck cases within this project.

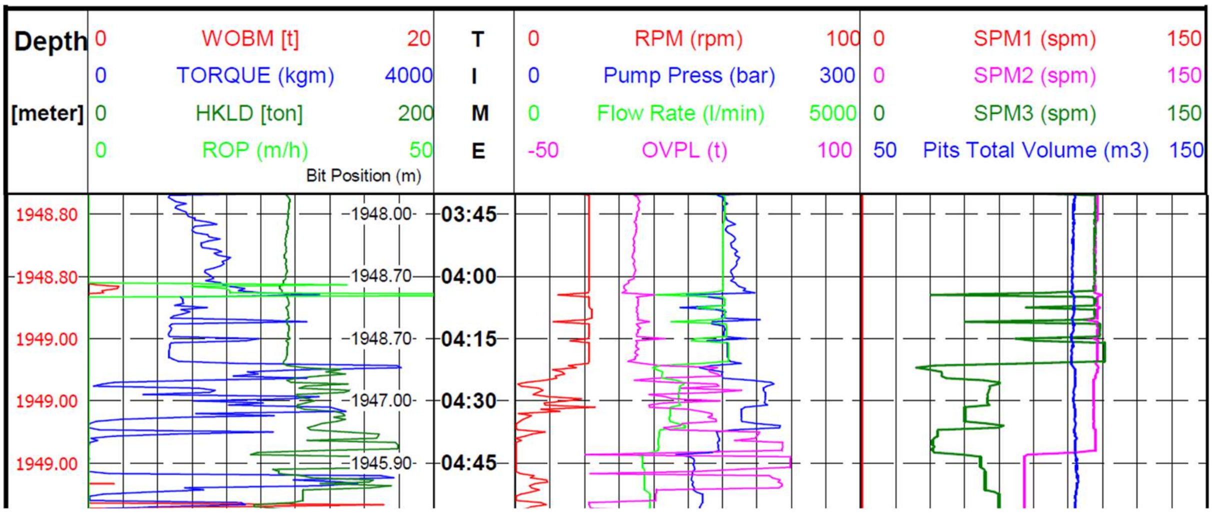


Figure 3-3 – Example of time log interval with pipe becoming stuck condition

3.2 Data preprocessing

Data preprocessing is one of the crucial steps for modeling using ML techniques, to ensure that produced model will provide accurate and reliable outputs, the results will not be biased and high predictive accuracy will be achieved. Data preprocessing typically starts from cleaning the data set from unnecessary noisy or duplicated data, finding missing data points, detecting outliers, understanding the distribution of data, followed by formatting and transforming the data (Hemmati-Sarapardeh et al., 2020).

3.2.1 Data processing software

Google Colaboratory (Google Colab) platform was used for this project to perform all the necessary actions for data preparation, processing and developing the intelligent models using Python codes. It is a cloud-based platform that allows the users to create and run Python codes for data modeling tasks, as well as gives an opportunity to visualize the output as it has access to Graphics Processing Units (GPU) and Tensor Processing Units (TPU). The platform has similar environment as Jupiter Notebook without having special hardware, with access to the main ML libraries, giving an opportunity to perform complex tasks of this research by evaluation of several ML techniques and comparison of results.

3.2.2 Dataset construction

Two ways of data construction was evaluated for this project: time-based and depth-based approaches. Depth-based approach brings significant challenges when dealing with recorded historical data, as depth-based data usually recorded once when the interval is drilled, hence, in case of additional trips or some off-bottom activities other than drilling, usually the depth-based surface data is not recorded and stored, unless there are special logging trips (e.g., wireline logging or LWD). Hence, it was decided to use time-based mud logging data, as it gets recorded and stored regardless of activity. Depth-based approach can be used for defining the issues only while drilling when the bit is on bottom, or in case if there is a real-time data streaming and storage is available with opportunity to separate time-index into depth-based data (Salminen et al., 2016).

Another challenge was to define the interval and number of data points for each case to make sure that the main data pattern is recorded, at the same time data set is not overloaded by unnecessary data. After some analysis of each case, it was decided to keep each data set with 10 representative data samples over certain interval of time that captures the stuck pipe incidents, which is equivalent to the drilling or reaming/tripping of one pipe stand which is approximately 28m long. As result, data base with 61 data sets, each data set containing 10 data points were constructed for this project.

3.2.3 Detection of missing data

Once data collection process was finalized the data preprocessing has been started from identification of missing data, null or duplicated values. It was found that main part of missing data points is coming from the data sources with update rates less than main mud logging time-based surface data. These data include drilling fluid information, data from daily drilling reports or final reports with higher update rates. Some data was absent from several wells, which lead to look for alternative sources to extract required information. Since some information is available once for the duration of one case, it was decided to fill in missing data with only one value applicable for this dataset, such as drilling mud information, lithology. Input data form is presented on the example of several datasets in the Table 3-5.

Table 3-5 – Input data set sample.

# of cases	Drilling Paramaters																	Mud Properties							Stuck type
	DPTS	HS	OPERATION	BL	ROP	WOBM	RPM	TORM	HKLA	PSPM	MFIN	SALT	CLAY	ANHYDRITE	SANDSTONE	LIME	OTHER	MW	MFV	PV	YP	GEL10SEC	GEL10MIN	SOLID	
	m	in		ft	min/m	ton	rpm	kgm	tons	atm	l/min	%	%	%	%	%	%	kg/dm3	sec/L	cp	cp	lbf/100ft^2	lbf/100ft^2	%	
1	1455.00	16.00	Drilling	351.64	3.38	9.39	31.95	1633.14	73.56	197.75	3352.11	5.00	0.00	95.00	0.00	0.00	0.00	1.26	44	13	11	3	5	2	NS
1	1455.00	16.00	Drilling	351.64	3.38	9.10	31.95	1739.64	74.15	195.21	3338.03	5.00	0.00	95.00	0.00	0.00	0.00	1.26	44	13	11	3	5	2	
1	1455.00	16.00	Drilling	351.64	4.19	8.73	31.95	1857.99	73.41	198.59	3352.11	5.00	0.00	95.00	0.00	0.00	0.00	1.26	44	13	11	3	5	2	
1	1455.00	16.00	Drilling	351.64	4.19	7.99	31.95	1786.98	73.41	198.59	3352.11	5.00	0.00	95.00	0.00	0.00	0.00	1.26	44	13	11	3	5	2	
1	1455.00	16.00	Drilling	351.64	3.26	8.36	31.36	1846.15	72.96	198.59	3352.11	5.00	0.00	95.00	0.00	0.00	0.00	1.26	44	13	11	3	5	2	
1	1455.00	16.00	Drilling	351.64	3.13	8.36	31.36	1834.32	72.81	196.06	3352.11	5.00	0.00	95.00	0.00	0.00	0.00	1.26	44	13	11	3	5	2	
1	1455.00	16.00	Drilling	351.64	2.74	7.62	31.95	1763.31	73.26	196.06	3366.20	5.00	0.00	95.00	0.00	0.00	0.00	1.26	44	13	11	3	5	2	
1	1455.00	16.00	Drilling	351.64	3.45	7.47	31.36	1597.63	72.81	194.37	3338.03	5.00	0.00	95.00	0.00	0.00	0.00	1.26	44	13	11	3	5	2	
1	1455.00	16.00	Drilling	351.64	4.11	7.25	31.36	1550.30	72.81	192.68	3352.11	5.00	0.00	95.00	0.00	0.00	0.00	1.26	44	13	11	3	5	2	
1	1455.00	16.00	Drilling	351.64	4.11	6.66	31.36	1550.30	72.67	191.83	3352.11	5.00	0.00	95.00	0.00	0.00	0.00	1.26	44	13	11	3	5	2	
2	1948.00	16.00	Drilling	351.64	0.00	0.00	20.00	1046.51	115.56	174.66	2524.00	95.00	0.00	5.00	0.00	0.00	0.00	1.3	43	12	12	3	4	4	S
2	1948.00	16.00	Drilling	351.64	5.42	4.51	21.00	1379.84	108.54	180.46	2640.00	95.00	0.00	5.00	0.00	0.00	0.00	1.3	43	12	12	3	4	4	
2	1948.00	16.00	Drilling	351.64	7.70	8.29	19.00	1434.11	103.89	188.59	2640.00	95.00	0.00	5.00	0.00	0.00	0.00	1.3	43	12	12	3	4	4	
2	1948.00	16.00	Drilling	351.64	24.58	11.05	19.26	2511.41	101.95	207.74	3010.00	95.00	0.00	5.00	0.00	0.00	0.00	1.3	43	12	12	3	4	4	
2	1948.00	16.00	Drilling	351.64	0.00	0.00	21.00	1007.75	114.40	194.97	2524.00	95.00	0.00	5.00	0.00	0.00	0.00	1.3	43	12	12	3	4	4	
2	1948.00	16.00	Drilling	351.64	5.72	1.74	16.53	2651.16	111.28	205.42	3000.00	95.00	0.00	5.00	0.00	0.00	0.00	1.3	43	12	12	3	4	4	
2	1948.00	16.00	Drilling	351.64	0.00	0.00	21.00	1147.29	114.73	179.89	2040.00	95.00	0.00	5.00	0.00	0.00	0.00	1.3	43	12	12	3	4	4	
2	1948.00	16.00	Drilling	351.64	33.08	17.04	12.45	1581.39	96.50	210.64	3010.00	95.00	0.00	5.00	0.00	0.00	0.00	1.3	43	12	12	3	4	4	
2	1948.00	16.00	Drilling	351.64	0.00	0.00	21.00	1245.73	119.84	181.62	2273.00	95.00	0.00	5.00	0.00	0.00	0.00	1.3	43	12	12	3	4	4	
2	1948.00	16.00	Drilling	351.64	6.55	1.51	11.09	2666.67	75.48	201.35	3010.00	90.00	0.00	10.00	0.00	0.00	0.00	1.3	43	12	12	3	4	4	
3	2332.00	12.25	Drilling	300.03	7.23	5.19	116.70	4138.20	85.84	224.66	2769.84	90.00	0.00	10.00	0.00	0.00	0.00	1.51	53	22	25	5	8	11	NS
3	2332.00	12.25	Drilling	300.03	9.20	5.13	116.97	4078.80	83.46	224.65	2770.38	90.00	0.00	10.00	0.00	0.00	0.00	1.51	53	22	25	5	8	11	
3	2332.00	12.25	Drilling	300.03	10.83	4.80	117.43	4138.20	84.54	224.72	2770.28	90.00	0.00	10.00	0.00	0.00	0.00	1.51	53	22	25	5	8	11	
3	2332.00	12.25	Drilling	300.03	9.07	4.64	117.84	3861.00	83.68	224.41	2770.31	90.00	0.00	10.00	0.00	0.00	0.00	1.51	53	22	25	5	8	11	
3	2332.00	12.25	Drilling	300.03	10.30	4.74	117.72	3870.90	84.54	224.41	2769.36	90.00	0.00	10.00	0.00	0.00	0.00	1.51	53	22	25	5	8	11	
3	2332.00	12.25	Drilling	300.03	10.87	4.40	117.88	3841.20	83.68	224.41	2769.83	90.00	0.00	10.00	0.00	0.00	0.00	1.51	53	22	25	5	8	11	
3	2332.00	12.25	Drilling	300.03	14.40	3.47	122.23	3742.20	84.76	224.37	2770.05	90.00	0.00	10.00	0.00	0.00	0.00	1.51	53	22	25	5	8	11	
3	2332.00	12.25	Drilling	300.03	11.70	3.86	120.48	3771.90	84.97	223.85	2769.37	90.00	0.00	10.00	0.00	0.00	0.00	1.51	53	22	25	5	8	11	
3	2332.00	12.25	Drilling	300.03	11.90	3.54	118.05	3771.90	84.54	223.41	2769.14	90.00	0.00	10.00	0.00	0.00	0.00	1.51	53	22	25	5	8	11	
3	2332.00	12.25	Drilling	300.03	5.97	6.03	116.75	4306.50	85.19	223.41	2770.39	90.00	0.00	10.00	0.00	0.00	0.00	1.51	53	22	25	5	8	11	
4	1950.00	16.00	Drilling	351.64	6.92	8.87	24.00	1198.44	108.17	180.70	2953.00	95.00	0.00	5.00	0.00	0.00	0.00	1.3	44	12	12	3	4	4	NS
4	1950.00	16.00	Drilling	351.64	6.73	8.98	24.00	1182.88	108.20	180.12	2953.00	95.00	0.00	5.00	0.00	0.00	0.00	1.3	44	12	12	3	4	4	
4	1950.00	16.00	Drilling	351.64	8.82	9.61	24.00	1252.92	107.78	179.53	2962.00	95.00	0.00	5.00	0.00	0.00	0.00	1.3	44	12	12	3	4	4	
4	1950.00	16.00	Drilling	351.64	9.06	9.84	24.00	1315.18	107.00	179.50	2962.00	95.00	0.00	5.00	0.00	0.00	0.00	1.3	44	12	12	3	4	4	
4	1950.00	16.00	Drilling	351.64	9.16	9.88	24.00	1354.08	107.00	180.70	2953.00	95.00	0.00	5.00	0.00	0.00	0.00	1.3	44	12	12	3	4	4	
4	1950.00	16.00	Drilling	351.64	9.56	10.04	24.00	1322.96	106.23	180.70	2953.00	95.00	0.00	5.00	0.00	0.00	0.00	1.3	44	12	12	3	4	4	
4	1950.00	16.00	Drilling	351.64	9.05	9.61	24.00	1330.74	106.22	180.70	2953.00	95.00	0.00	5.00	0.00	0.00	0.00	1.3	44	12	12	3	4	4	
4	1950.00	16.00	Drilling	351.64	9.30	9.88	24.00	1346.30	106.22	180.12	2953.00	95.00	0.00	5.00	0.00	0.00	0.00	1.3	44	12	12	3	4	4	
4	1950.00	16.00	Drilling	351.64	9.30	9.80	24.00	1354.09	105.83	180.70	2953.00	95.00	0.00	5.00	0.00	0.00	0.00	1.3	44	12	12	3	4	4	
4	1950.00	16.00	Drilling	351.64	9.52	10.50	24.00	1392.00	105.00	181.20	2944.00	95.00	0.00	5.00	0.00	0.00	0.00	1.3	44	12	12	3	4	4	
5	1954.00	16.00	Drilling	351.64	25.25	9.92	24.30	1163.43	105.26	172.38	2914.00	95.00	0.00	5.00	0.00	0.00	0.00	1.31	44	12	12	3	4	4	S
5	1954.00	16.00	Drilling	351.64	5.28	0.94	25.10	986.15	114.20	169.06	2900.00	95.00	0.00	5.00	0.00	0.00	0.00	1.31	44	12	12	3	4	4	
5	1954.00	16.00	Drilling	351.64	32.00	9.86	25.40	1185.60	106.37	174.86	2925.00	95.00	0.00	5.00	0.00	0.00	0.00	1.31	44	12	12	3	4	4	
5	1954.00	16.00	Drilling	351.64	8.35	1.44	25.70	1008.31	111.36	179.83	2914.00	95.00	0.00	5.00	0.00	0.00	0.00	1.3	44	12	12	3	4	4	
5	1954.00	16.00	Drilling	351.64	10.94	6.76	23.50	1894.76	108.58	190.60	2872.00	95.00	0.00	5.00	0.00	0.00	0.00	1.3	44	12	12	3	4	4	
5	1954.00	16.00	Drilling	351.64	6.62	1.66	21.50	1130.19	114.68	178.18	2928.00	95.00	0.00	5.00	0.00	0.00	0.00	1.3	44	12	12	3	4	4	
5	1954.00	16.00	Drilling	351.64	11.24	6.93	25.80	2736.84	108.58	146.68	2458.00	95.00	0.00	5.00	0.00	0.00	0.00	1.3	44	12	12	3	4	4	
5	1954.00	16.00	Drilling	351.64	8.82	4.46	25.40	3013.81	112.82	178.35	2825.00	95.00	0.00	5.00	0.00	0.00	0.00	1.31	44	12	12	3	4	4	
5	1954.00	16.00	Drilling	351.64	6.73	1.14	14.60	954.64	117.16	131.10	2312.00	95.00	0.00	5.00	0.00	0.00	0.00	1.31	44	12	12	3	4	4	
5	1954.00	16.00	Drilling	351.64	10.86	5.21	26.20	2138.10	112.82	173.62	2864.00	95.00	0.00	5.00	0.00	0.00	0.00	1.31	44	12	12	3	4	4	

3.2.4 Detection of outliers and data distribution analysis

Identification of outliers and its analysis is at the highest importance of data preprocessing stage. Outliers usually refer to the data points at its extreme values. These outliers can significantly affect the model performance and its reliability, as ML algorithms learn data patterns. There are different ways for outliers' detection, including numerical approaches and graphical methods (Hemmati-Sarapardeh et al., 2020). In this research, graphical methods were used for data visualization capabilities which allows to perform data distribution analysis. This analysis is an important step of data preprocessing stage, that helps to understand if the outlier is an error measurement or real phenomenon. This is critical point for stuck pipe detection problem, as we try to define anomalies of data patterns to identify the early signs of stuck pipe condition.

Number of visualization techniques were used for outlier detection and data distribution analysis, such as box plot, swarm and boxen plot, violin plot, histogram and scatter plots. From the swarm and boxen plot, better distribution of data at the tails can be defined, especially for big set of data. From Figure 3-4 the following observations of the surface drilling data distribution can be made:

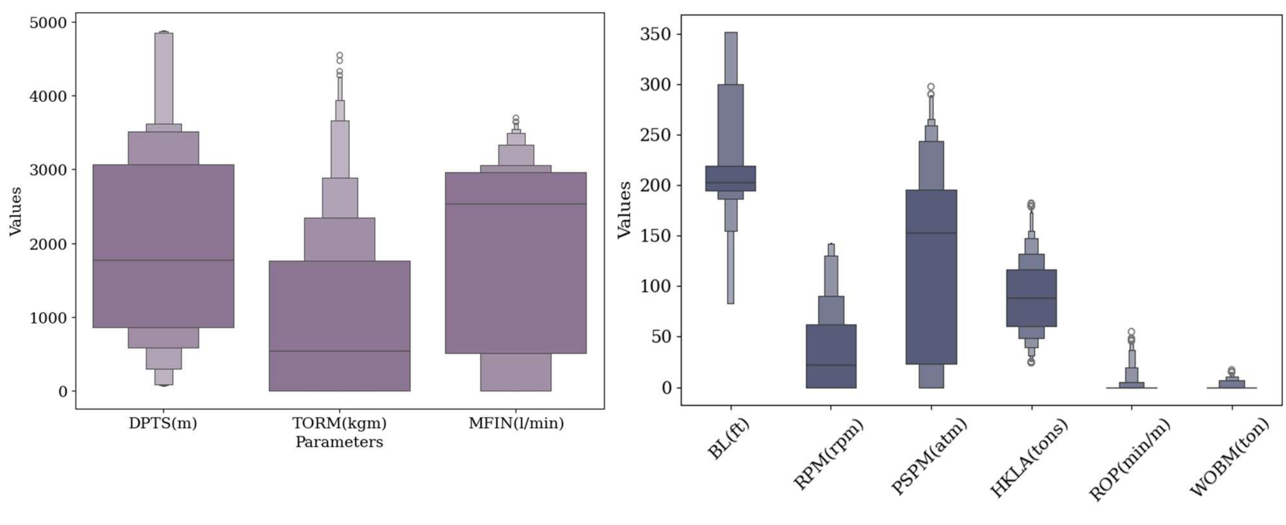


Figure 3-4 - Boxen plot of surface drilling data distribution

- DPTS – depth measurements significantly vary across big range of data, showing that stuck cases are seen at different intervals, however the main concentration of data is at depths of around 1000 m to 3000 m, which corresponds to massive salt formations.
- TRQM – torque values also demonstrate notable changes across big interval with concentration of main data points in the range of 0 to 2000 kgm, which is common behavior for stuck situation where torque significantly fluctuates over big range of data.
- RPM – repeats the torque behavior, as these two measurements have strong relations when the pipe is stuck because of inability to rotate the string.
- PSPM – circulation pressure has a big range of data, as it varies with the depth.

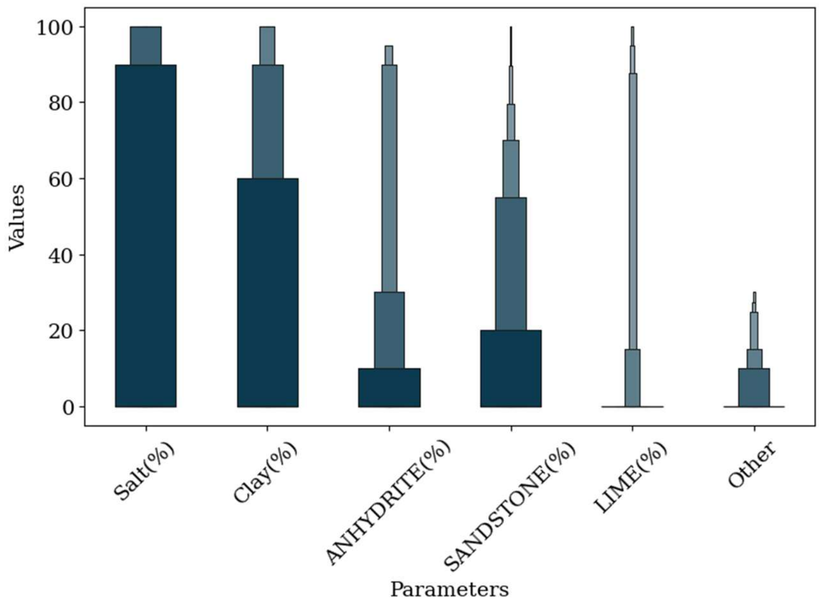


Figure 3-5 – Boxen plot of geological data distribution

From Figure 3-5 of geological data distribution, it is clearly seen that the main cases of dataset correspond to the presence of salt formation and clays, distribution of sandstone also demonstrates that the stuck cases occurred in the layered intervals.

Another approach using box and violin plot, which gives an opportunity to see detailed distribution for deeper analysis, with probability distribution represented by its outer shape based on Kernel Density Estimation (KDE).

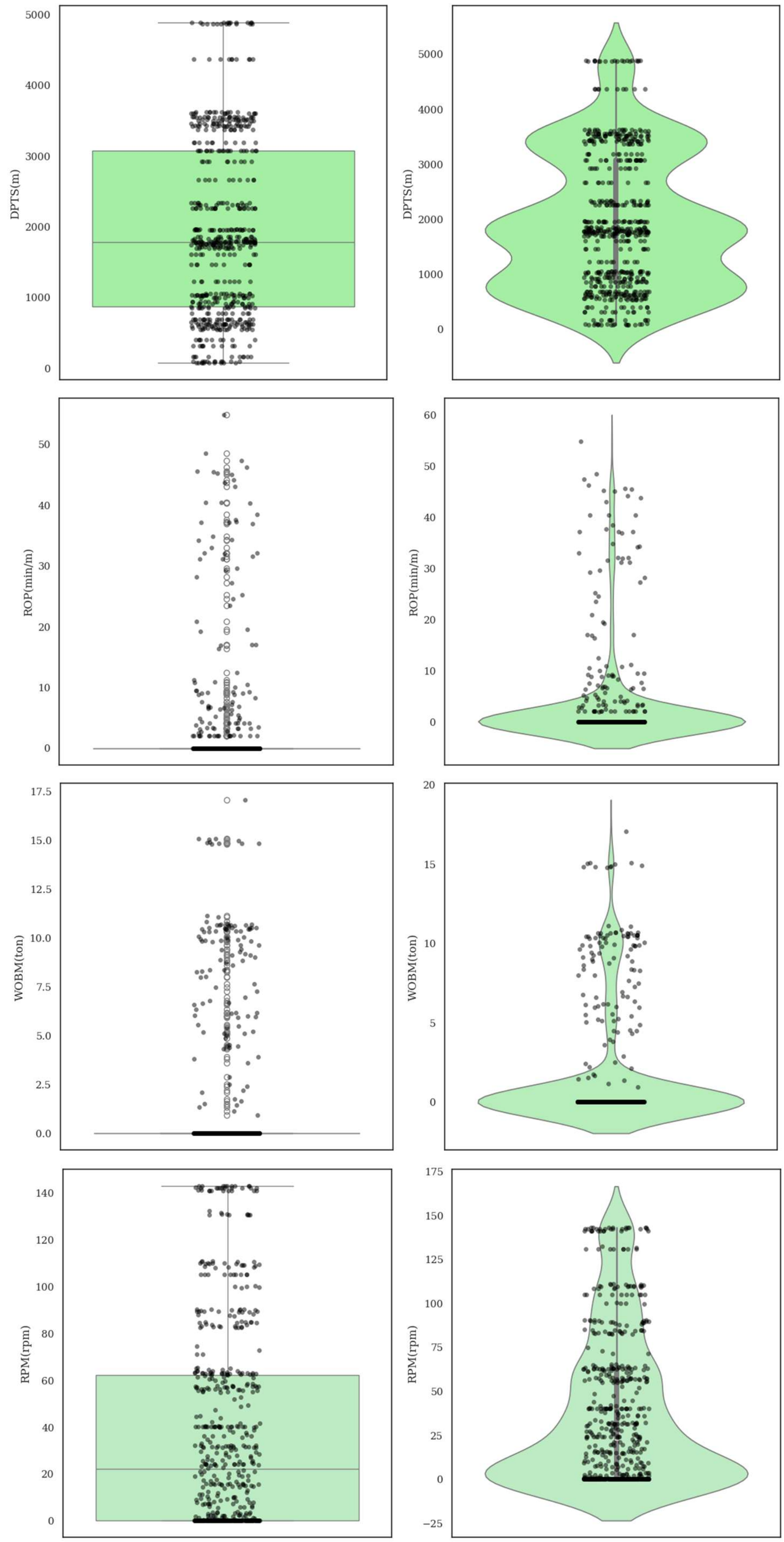


Figure 3-6 and Figure 3-7 show probability distribution of each parameter together with swarm data points. From these plots understanding of outliers' distribution can be studied in more details. As example, data points for torque readings outside of the main distribution is not because of error measurements, but real values that correspond to spiky behavior of torque during stuck condition.

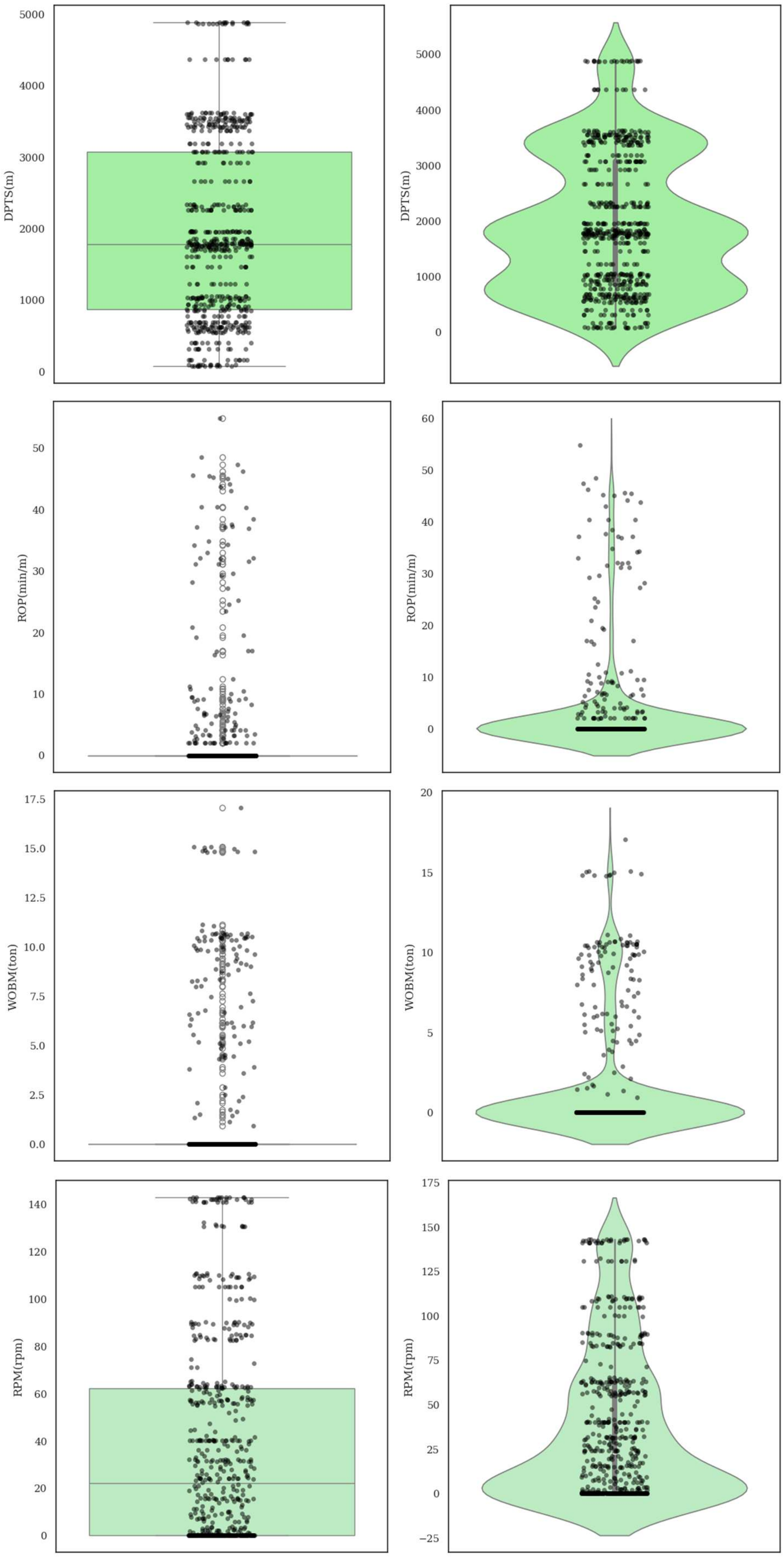


Figure 3-6 – Box and violin plot for surface drilling parameters.

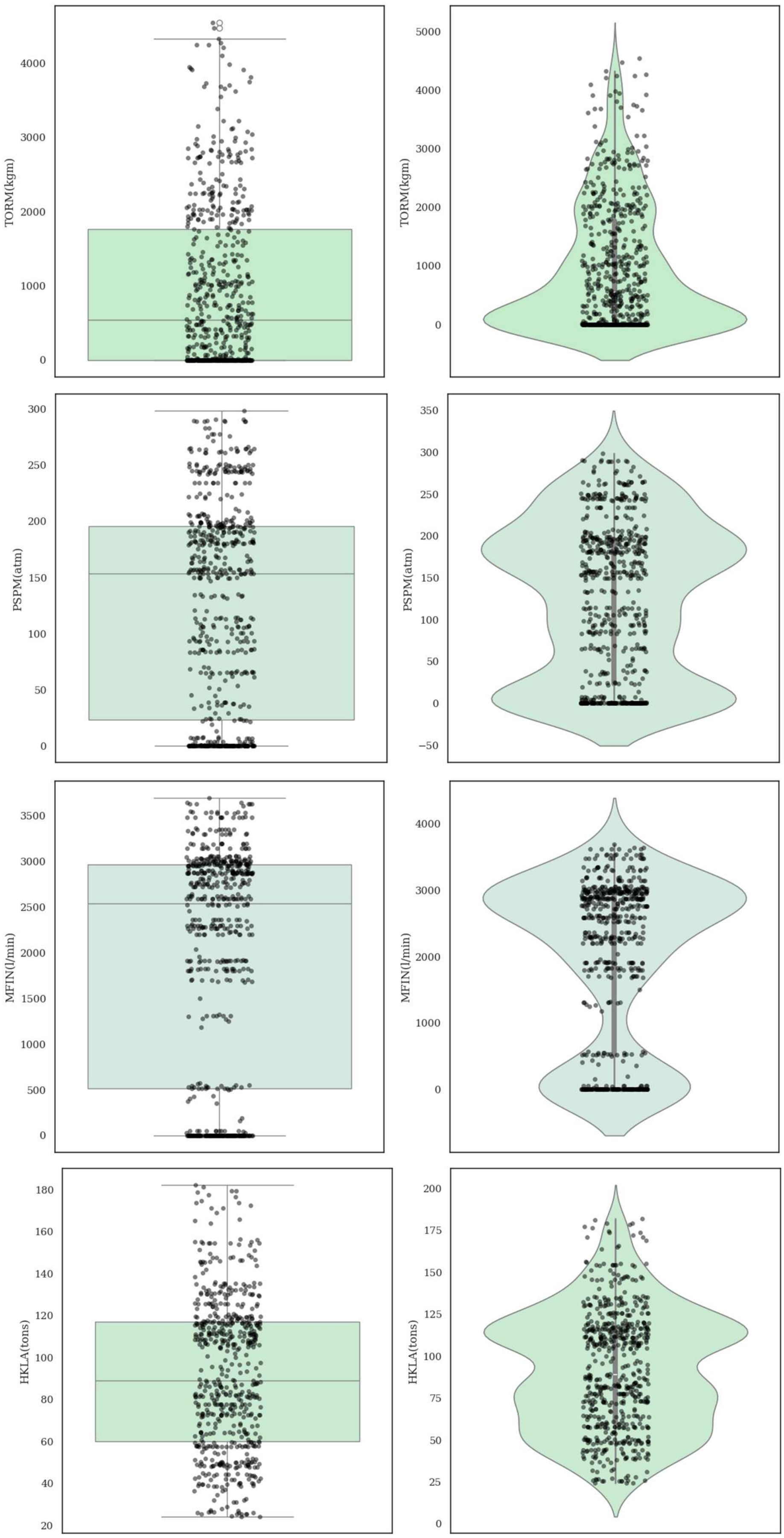


Figure 3-7 – Box and violin plot for surface drilling parameters (continued).

3.3 ML Algorithms

3.3.1 ML model selection

From extended literature review on the stuck pipe detection problem and analysis of existing ML techniques, it was decided to build and compare different intelligent models based on supervised ML methods to define the most reliable and accurate model for existing problem. From number of researches, as noted previously, supervised ML algorithms such as LR, KNN, SVM and DT have demonstrated one of the highest performances with low error measurements and false alarm detection rates (Jahanbakhshi et al., 2012; Chamkalani et al., 2013; Magana-Mora et al., 2019; Alshaikh et al., 2019; Ahmed et al., 2019; Elmousalami & Elaskary, 2020; Al Dushaishi et al., 2021; Brankovic et al., 2021; Mal et al., 2022). In most of the cases, other ML approaches performed better than ANN models (Albaiyat, 2012; Magana-Mora et al., 2019; Alshaikh et al., 2019). Hence, these algorithms were used in the current research. In addition, ensemble learning boosting algorithms that also demonstrated high accuracy in predicting stuck pipe incidents were also studied (Ahmed et al., 2019; Elmousalami & Elaskary, 2020; Brankovic et al., 2021; Malki et al., 2024). After extensive research, it was found that the algorithm that was not used in any of the previous studies of stuck pipe, however, has high potential for prediction, known as Categorical Boost algorithm was used to build predictive model for stuck pipe detection. The data modeling flow chart of this research is presented in Figure 3-8. Each of the smart models will be discussed in more details further in this chapter.

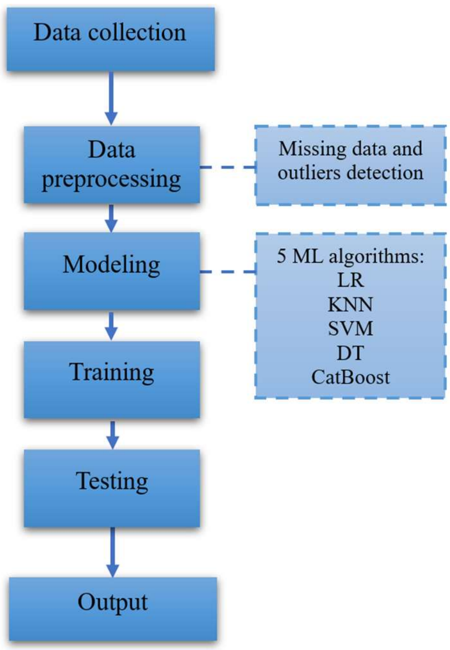


Figure 3-8 – Data processing and model development flow chart

3.3.2 Logistic Regression algorithm

The main algorithm behind this model is known as Sigmoid or Logistic function, which identifies the relationship between dependent and independent variables, is shown in

Equation 3-1 below (Elmousalami & Elaskary, 2020):

$$P = \frac{1}{1+e^{-(a+bX)}} \qquad \text{Equation 3-1}$$

Where:

P – probability of classification

e – natural logarithm base (Euler's number)

a and b – coefficients of the model

The output of the model ranges from 0 to 1, and can be graphically represented in S shaped curve. It gives smooth output that used for prediction probability. If the output is higher than predefined threshold, it will be classified as 1, otherwise will be classified as 0. Similarly to the other types of supervised ML algorithms, it requires the data for training and testing the model. At the training phase the model defines the coefficients of a and b to achieve the lowest classification error. It is done with the help of Maximum Likelihood Estimation (MLE) method. Unlike the Linear Regression which uses Ordinary Least Squares (OLS) method, LR requires non-linear method to reduce the classification error. This ML method is one the simplest but efficient classification approaches, however, it struggles when the data sampling rate is not consistent and there is high correlation between variables, known as multicollinearity. In addition, this ML method is sensitive to outliers, hence, preprocessing step will be critical when dealing with this method (Elmousalami & Elaskary, 2020).

3.3.3 K-nearest neighbors (KNN) algorithm

KNN which is known as instance-based learning algorithm evaluates new sample data points through determining the class label from their closest k neighboring instances. The algorithm proceeds as follows: firstly, set the fixed value of K (the number of nearest neighbors). Then the distance calculation between a new data point and every training data example must be computed (using Euclidean distance as an example), followed by selecting the K through evaluating their shortest distances (Singh et al., 2013). And finally, new data point receives its class label by using the most common class from its k-nearest neighbors.

The KNN classifier operates as follows: for a test observation x_0 and a predefined integer K, the algorithm identifies the K nearest training points (denoted N_0) to x_0. The conditional probability for class j is computed as the proportion of points in N_0 belonging to class j as shown in Equation 3-2 (James et al., 2021):

$$P_r(Y = j|X = x_0) = \frac{1}{K}\sum_{i \in N_0} I(y_{i=j}) \quad \text{Equation 3-2}$$

Where, $I(y_{i=j})$ is an indicator function. As the result, the classifier then uses Bayes' rule to allocate x to the highest estimated probability class. It follows that value of K has a significant effect on the right classification of a new data point. Hence, to reduce the misclassification probability, the test data point should be assigned to the class with the highest posterior probability. Meaning that, the K-nearest neighbors (KNN) algorithm identifies the K nearest points within the training dataset and then assigns the new data point to the most commonly present class type in this discovered proximity. Very small number of K may lead in creation of many classes, thus being sensitive to outliers, in opposite, very big number of K may lead to misclassification of data because of overfitting issues (Bishop, 2006). In the example illustrated in Figure 3-9, data set is run with different numbers of K, the first case when K = 1 and the second one when K = 100. None of the given examples show good prediction performance. When K = 1 the decision boundary is extremely flexible demonstrating very high variations. In the case of K = 100, the boundary flexibility notably dropped becoming more linear, which also does not correspond to the Bayes decision boundary curve, thus misclassifying the data points. Hence, optimal number of K is critical for this algorithm (James et al., 2021).

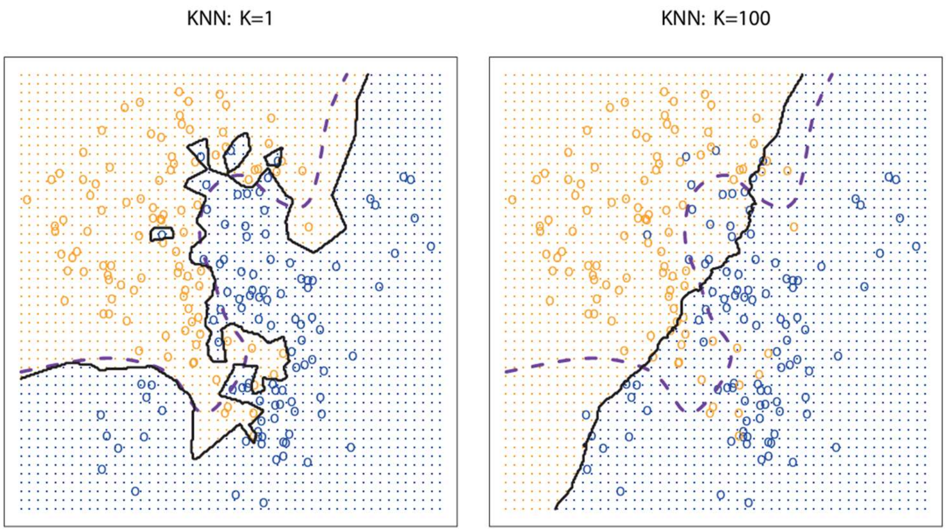


Figure 3-9 – Comparison plots of K=1 to K=100 data sets, where solid black curves correspond to the KNN output and purple dushed line to the Bayes decision boundary (James et al., 2021).

3.3.4 Support Vector Machine (SVM) algorithm

The SVM algorithm starts from defining the optimal hyperplane that can increase the margin between two groups of data. The data points closest to the hyperplane is taken into account to identify the final decision boundary, these points are known as support vectors. And finally, increasing the margin, the distance from hyperplane and closest support vector, notably enhances the generalization of the model.

The simplest SVM is Maximum Margin Classifier (MMC), this model is suitable for linearly separable data. The problem is framed to minimize the size of the boundary's parameters while ensuring all data points are correctly classified with a fixed functional margin (set to 1 for simplicity). The geometric margin (actual distance from the boundary to support vectors) is maximized by adjusting the boundary's orientation. The geometric margin for the case can be shown in Equation 3-3 and Equation 3-4 below (Cristianini & Shawe-Taylor, 2000):

$$\langle w \cdot x^{+} \rangle + b = +1 \quad \text{Equation 3-3}$$

$$\langle w \cdot x^{-} \rangle + b = -1 \quad \text{Equation 3-4}$$

Where:

w – is a weight vector

x^{+} – positive feature vector

x^{-} – negative feature vector

b – bias term

The optimization objective, known as optimal design of the classification algorithm of hyperplane with certain margin (Abbas et al., 2019), for the given case become as follows, resulting in the maximal margin hyperplane with geometric margin y (Equation 3-5):

$$\min_{w,b} \ \frac{1}{2} \|w\|^{2} \quad \text{Equation 3-5}$$

Subjected to:

$$y_i(\langle w \cdot x_i \rangle + b) \geq 1,\ i = 1, \ldots, l.$$

Where:

$\|w\|_2$ – margin maximization parameter, the distance between two hyperplanes becomes $\frac{2}{\|w\|_2}$.

The maximal margin hyperplane illustrated in the Figure 3-10, where the points closest to the hyperplane known as support vectors are also highlighted.

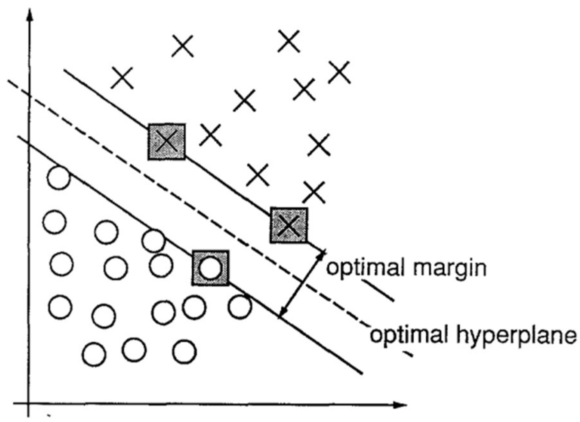


Figure 3-10 – The illustration of maximal margin hyperplane with support vectors (Vapnik & Cortes, 1995).

As noted previously, maximal margin classifier is suited to solve linear problems. However, for more complex overlapping cases or data with high noise and presence of outliers, soft margines should be used to address the real-life cases. The mechanism works by allowing to misclassify some data points to increase the margin and allowing for better overall classification. The algorithm works by introducing non-negative variable ξ_i, showing the degree of misclassification as follows (Equation 3-6) (Cristianini & Shawe-Taylor, 2000):

$$y_i(\langle w \cdot x_i \rangle + b \geq 1 - \xi_i, \qquad \text{Equation 3-6}$$

$$i = 1, \ldots, l, \ \xi_i \geq 0,$$

The following optimization objective will become as follows (Equation 3-7):

$$\min_{w,b,\xi} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l} \xi_i \qquad \text{Equation 3-7}$$

Where, C – parameter that regulates the optimal choice of margin size and classification error that varies withing wide range of values. Defining the optimal value of C critical for the classification problem, as if it's too large, misclassification can be reduced but resulting in overfitting, and opposite, if it's too small, it may lead to high degree of misclassification, consequently, increasing generalization.

$\sum_{i=1}^{l} \xi_i$ – misclassification penalty

Another way of handling non-linear cases is by use of a Kernel trick. In the Kernel method data transforms into high-dimensional spaces called feature spaces through which hyperplane separation becomes achievable (Rostami & Manshad, 2014). The methodology of Kernel functions works through mapping input set of x parameters into a new space of $\phi(\boldsymbol{x})$ as follows (Equation 3-8) (Cristianini & Shawe-Taylor, 2000):

$$\mathbf{x} = (x_1, \dots, x_n) \rightarrow \phi(\boldsymbol{x}) = (\phi_1(x), \dots, \phi_N(x)) \qquad \text{Equation 3-8}$$

The Figure 3-11 illustrates a feature mapping, where initially input data cannot be linearly separated into two groups, while feature mapping helped to change the representation of data and classification become achievable.

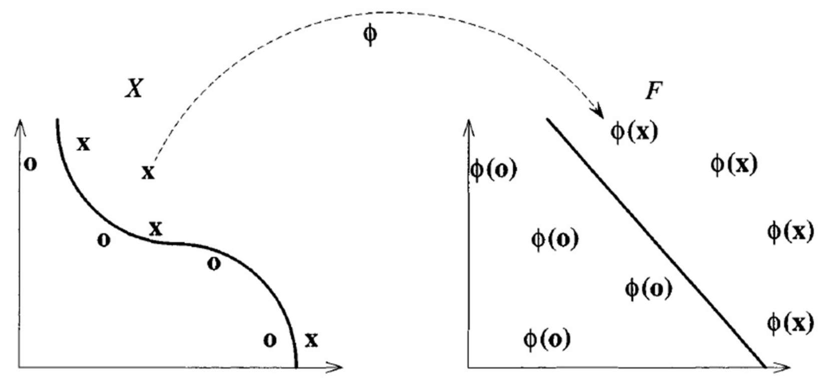


Figure 3-11 – Feature mapping example of data classification using Kernel function (Cristianini & Shawe-Taylor, 2000).

3.3.5 Decision Tree algorithm

Another powerful tool used in data mining for classification and regression problems is Decision Trees. Decision Trees work based on classification and regression trees (CART) algorithm, it's a basic tree-based model, known for its simplicity in interpretation and performance, is used to build the smart model (Breiman et al., 2017). It divides the feature space into rectangles and runs a model in each of them. A Decision Tree contains three essential elements including its starting point which handles the entire dataset while picking the best feature for splitting. Decision points in the data function as internal nodes which divide information according to feature threshold values. The last stage of the Decision Tree consists of Leaf Nodes which present the ultimate prediction outcome. The decision paths appear through branches that link nodes together. Example of CART model is shown in the Figure 16, where initially the space was divided into two areas and the Y response through local mean values in these areas are determined. Selection of the variable and split-point allows to achieve the most accurate model. The total space repeats splitting into two regions which continue until a predefined stopping rule is used (Hastie et al., 2009). The output of the model are the leaves that corresponds to the regions R_1 to R_5. Figure 3-12 shows the binary output of this splitting in the form of the growing tree.

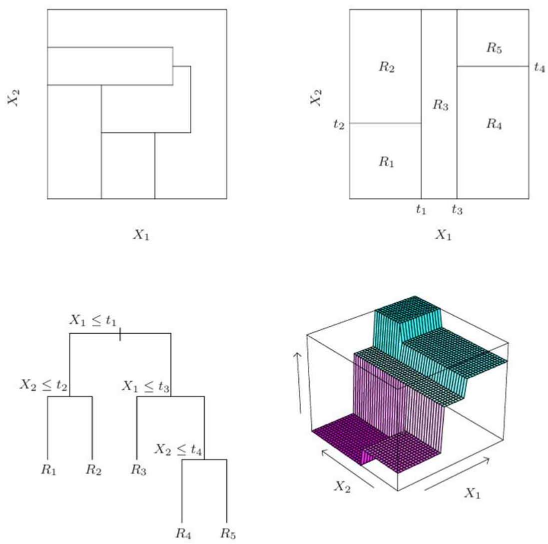


Figure 3-12 – Example of CART classification (Hastie et al., 2009).

DT algorithm makes a decision for splitting based on impurity measurements, such as Gini Index (Equation 3-9) and Cross-entropy (Equation 3-10) (Bourel & Segura, 2018):

Gini Index: $G(S) = \sum_{k=1}^{K} p_K(1 - p_K)$ Equation 3-9

Where, p_k – proportion of class k in node S, K – number of classes, Cross-entropy of parent node: $H(S) = -\sum_{k=1}^{K} p_k log\ p_k$ Equation 3-10

Despite the number of advantages, DT has notable drawback that mainly related to accuracy of output data, as any changes in the input data can significantly affect the results of decision tree. In this regard, ensemble learning algorithms can help to overcome this issue by combining several ML approaches (Hastie et al., 2009). One of this method that was used in this project will be discussed in the next section.

3.3.6 CatBoost – gradient boosting algorithm

CatBoost or Categorical Boosting is advanced gradient boosting approaches that was developed to work with categorical parameters. The algorithm that was recently developed is known for its ability to work with highly diverse input parameters, handing outliers and noisy data and other complex environments. CatBoost operates through gradient boosting decision trees (GBDT) modified by several enhancements that affect its performance. CatBoost uses symmetric (oblivious) trees to split data at the same feature table along with threshold value within every node of the tree structure. Oblivious Trees provide two benefits including accelerated training and inference and reduced model variance that minimizes overfitting risk. The traditional gradient boosting techniques allow target leakage to occur during residual computation processes. CatBoost uses ordered boosting to split data into multiple permutations and it ensures the target information remains out of reach until an instance completes training (Dorogush et al., 2018).

When handling categorical data CatBoost uses categorical encoding approach as its data processing method: target-based encoding with ordered statistics using target statistics. CatBoost changes categorical attributes into numerical properties instead of applying the standard one-hot coding technique. The permutations method ensures data leakage prevention through this process. A categorical feature with less than 255 unique values are encoded with one-hot encoding (Dorogush et al., 2018).

As it seen from the Figure 3-13, the algorithm starts with an initial basic model that uses single leaf node to assign the overall dataset average value. Errors will be detected through larger circles and new decision trees are incorporated using gradient boosting approach. A new tree is trained during each iteration with the goal of minimizing mistakes from the preceding tree (Vaferi et al., 2024). The ensemble predictions become more refined as it combines output results from every tree. The error rate decreases with every new tree included. The process ends after satisfying pre-established criteria which include a minimum number of samples per split or a maximum depth for the trees (Yousefzadeh et al., 2024).

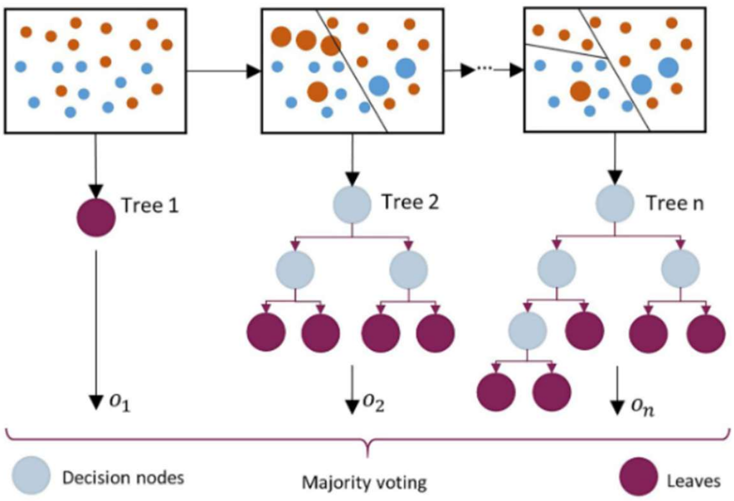


Figure 3-13 – CatBoost architecture (Yousefzadeh et al., 2024).

3.4 Training and testing of the developed models

Each intelligent model was trained and tested based on input dataset. The training phase also known as a learning phase is an important step, where the model learns data patterns from the input information and main parameters of the model get adjusted to achieve required outputs. As soon as the model passes the training stage, it can be then further run through testing phase, where the tuned model goes though unseen data to recognize the new data points and classify them based on previously learned patterns. Capability to accurately classify new data is very crucial for each ML-based model, and known as a generalization ability (Bishop, 2006).

3.4.1 Data splitting

In this research, because of diversity of input parameters and imbalanced data related to different parameters, it was decided to use stratified shuffle splitting technique to make sure that proper data distribution in the training and testing phases is achieved. This type of splitting in comparison to the random one is helping to avoid biased outputs, improve the overall model performance for the problem of stuck pipe detection, as it is a multi-class problem with limited data set samples. This technique randomly shuffles the data before dividing it to training and testing sets. Different number of splits and proportions of training and testing the data were used before the best proportion was found as 3:2: Number of splits: 1, test size: 60% for training, and 40% for testing.

3.5 Model performance evaluation metrices

Classification algorithms can be evaluated based on several parameters that demonstrate overall model performance, its accuracy, predictive capabilities and reliability. Accuracy measure standalone cannot demonstrate the real performance of the classifiers, especially when imbalanced data is used, hence, there are number of parameters should be considered when evaluating intelligent models (Mal et al., 2022). In this project, the following parameters were used (Table 3-6):

Table 3-6 – Intelligent models evaluation metrices:

Measure	***Formula***	***Definition***
Accuracy	$\frac{TP + TN}{TP + TN + FP + FN}$	The ratio of correctly classified instances between total instances including positives and negatives
Precision	$\frac{TP}{TP + FP}$	Ratio of accurate positive predictions to all the predictions marked as positive
Recall	$\frac{TP}{TP + FN}$	The percentage of actual positive cases that a model correctly identifies as positive instances
F1 score	$\frac{2 \times Precision\, X\, Recall}{Precision + Recall}$	A combination of precision and recall to produce a single evaluation measure

Where, TP – True Positive, is prediction of stuck situation corresponds with the real stuck case, FP – False Positive, is prediction of stuck situation when there is no actual stuck case, TN – True Negative, is prediction of non-stuck case corresponds to the real non-stuck case, and FN – False Negative, is prediction of non-stuck situation, when there is stuck condition

These parameters are calculated based on confusion matrix, a special tool that shows predicted label results compared to actual labels in a table structure for the classification problems (Othman et al., 2022). Confusion matrix delivers specific information about model errors that assists decision-making process, especially when false positive and false negative errors result in varying impacts. Confusion matrix for binary classification illustrated in the Figure 3-14, that can be also generated for multi-classification problems.

		Predicted Class		
		Positive	**Negative**	
Actual Class	**Positive**	True Positive (TP)	False Negative (FN)	**Sensitivity** $\frac{TP}{TP+FN}$
	Negative	False Positive (FP)	True Negative (TN)	**Specificity** $\frac{TN}{TN+FP}$
		Precision $\frac{TP}{TP+FP}$	**Negative Predictive Value** $\frac{TN}{TN+FN}$	**Accuracy** $\frac{TP+TN}{TP+TN+FP+FN}$

Figure 3-14 – Confusion matrix for binary classification task

Another metrics used is ROC (Receiver Operating Characteristics) graph. ROC operates as a performance evaluation method used to address accuracy metric limitations when dealing with unbalanced data and varying complex framework structures. TP rate is plotted on the Y axis, while FP rate appears on the X axis in ROC graph. The ROC graph illustrates how different trade-offs between true positives and false positives relate to each other. The scalar metric AUC (Area Under Curve) represents performance level from 0.5 to 1.0 and establishes higher ratings for superior classifiers (Fawcett, 2006).

3.5.1 K-fold cross validation

In addition, k-fold cross-validation technique was utilized in order to assess the model performance and mitigate the overfitting issues. In the beginning, the dataset is divided into k equal folds for maintaining the class balance in imbalanced datasets to achieve unbiased validation outcomes. The method conducts each iteration by allocating k−1 folds into training data, and applying remaining single fold to validate classification metrics, such as accuracy and F1-score resulting in one-time validation of all dataset points (Figure 3-15). The process ends by averaging k performance metrics from all folds, so the model can achieve a finalized generalization ability assessment that reduces the variance from separate train-test splits.

Detection of overfitting becomes possible with this technique, because it demonstrates inconsistent validation score distributions across folds, which indicates the model's sensitivity and potential instability (Alshaikh et al., 2019). The method maintains ideal performance metrics and class distribution quality while running efficiently which makes it essential for development of reliable classification evaluation systems. One of the advantages of this technique is that this evaluation method delivers optimal data usage for training and evaluation purposes on small datasets (Elmousalami & Elaskary, 2020). 5-fold cross validation technique was applied to each model for performance evaluation.

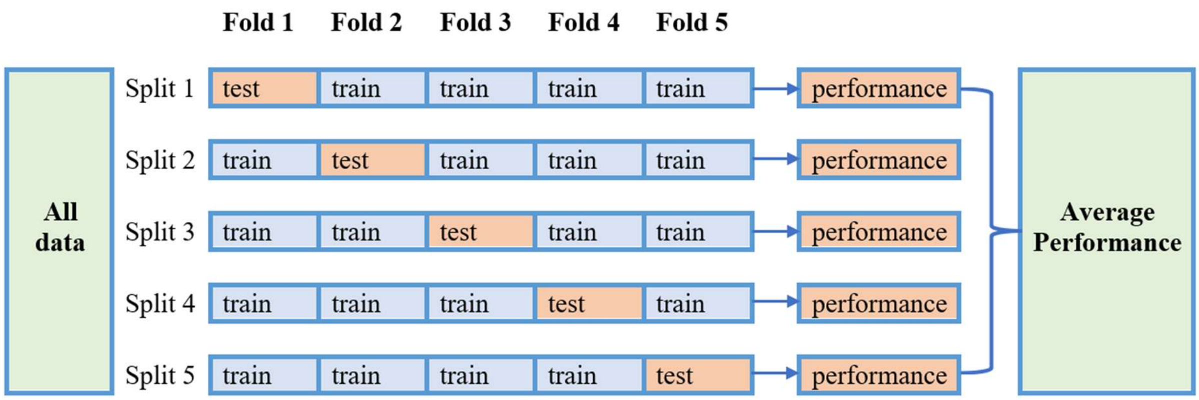


Figure 3-15 – 5-fold cross validation algorithm used for model performance evaluation.

4. RESULTS AND DISCUSSION

In this chapter, the results of stuck pipe modeling using ML approaches are discussed in more details. The first part includes data accuracy and error calculations of each developed model and their performance evaluation. It is followed by results validation section. In the next part, a detailed sensitivity analysis conducted to define the feature importance parameters of the most effective intelligent model is presented. The chapter is concluded with field implications of developed intelligent models.

4.1 Data accuracy and error calculations

As highlighted in previous chapters, five ML methods were used to develop efficient and reliable stuck pipe prediction methodology based on the given field data. The models have been applied to each hole section separately to evaluate their performance in different environment where stuck pipe cases were detected, which include 16in, 12.25in and 8in sections drilled, and the final evaluation is performed based on whole field data. 26 in section was eliminated from individual review as only one stuck pipe incident was detected while drilling this interval. However, this stuck case was included into the entire field evaluation. The models were tested to validate prediction capabilities with 40% of dataset. The smart models were used to define five types of stuck condition presented below:

- S – stuck pipe
- TS – tight spot
- PO – pack off
- OP – overpull
- NS – non-stuck

The final evaluation was performed on entire field data. Data accuracy and error estimation started from confusion matrix analysis of each model for the given hole section and entire field. The error estimations and accuracy calculations are presented in the Table 4-1:

Table 4-1 – Data accuracy and error calculations for each method:

Section	Measure	LR	KNN	SVM	DT	CatBoost
16in	Accuracy	90.62%	70.83%	62.5%	98.96%	100%
	Precision	0.91	0.73	0.67	0.99	1
	Recall	0.91	0.71	0.62	0.99	1
	F1 score	0.90	0.71	0.60	0.99	1
	AUC	0.92	0.90	0.92	0.98	1
12.25in	Accuracy	95%	89.17%	58.33%	96.67%	99.17%
	Precision	0.95	0.89	0.48	0.97	0.99
	Recall	0.95	0.89	0.58	0.97	0.99
	F1 score	0.95	0.89	0.50	0.97	0.99
	AUC	0.94	0.95	0.80	0.97	1
8in	Accuracy	87.50%	68.75%	68.75%	100%	100%
	Precision	0.90	0.81	0.72	1	1
	Recall	0.88	0.69	0.69	1	1
	F1 score	0.87	0.65	0.68	1	1
	AUC	0.94	1	0.25	1	1
Field	Accuracy	84.02%	81.56%	58.20%	99.59%	98.77%
	Precision	0.84	0.83	0.52	1	0.99
	Recall	0.84	0.82	0.58	1	0.99
	F1 score	0.84	0.81	0.54	1	0.99
	AUC	0.95	0.95	0.77	1	1

From the model evaluation metrices the following observations were made based on their application:

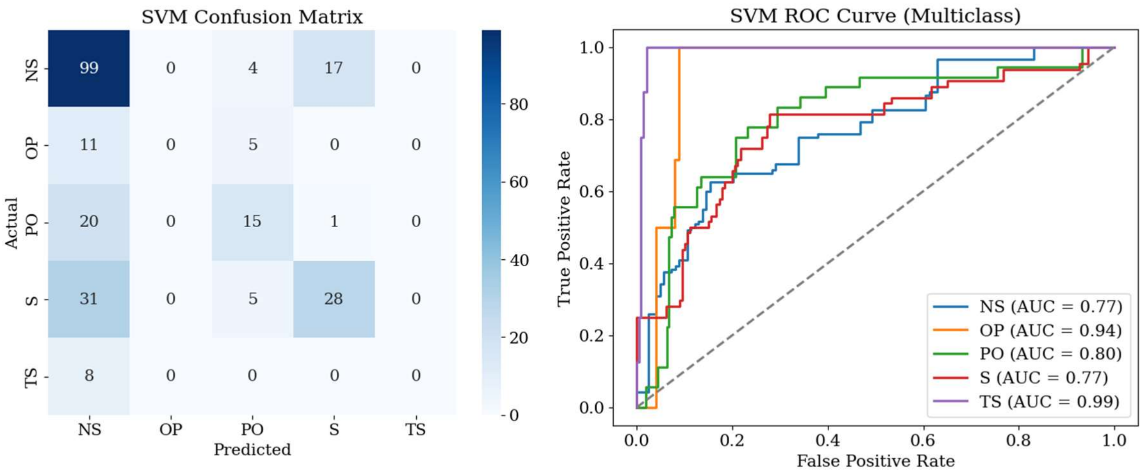
SVM method

The model exhibited significantly inferior performance relative to other approaches, achieving the maximum accuracy of 68.75% in the 8in section, while the lowest accuracy recorded in the entire field application was 58.2%. The highest performance of the model observed in the 8in section, where the correctly identified positive cases had a precision value of 0.72. The recall of real positive instances is significantly low, reaching a maximum value of 0.69. Important evaluation of F1 score is notably low, ranging from 0.5 to 0.68.

The overall balance between correctly and incorrectly detected positive cases can be seen from ROC curve in the (b)

Figure 4-1 – (a) SVM confusion matrix and (b) SVM ROC curve for entire field. (b). The model successfully detects both tight spots (TS) and overpulls (OP) following the evaluation of AUC for each stuck category, while showing a weakness in identifying non-stuck (NS) and stuck (S) samples.

Although SVM model is known as reliable classifier, its application for the specified dataset demonstrated lower results relative to other models. One of the main reasons of such evaluation can be because of its mechanism, as previously mentioned, SVM encounters difficulties with multi-classification problems and data imbalance, necessitating the application of additional preprocessing procedures. Furthermore, because of its sensitivity for outliers and noisy data, some datapoints may be classified by the model as noisy because of significant parameter fluctuations (e.g., torque values). The model requires hyperparameter tuning to attain improved outcomes in each case, which make this method time-extensive.



(a) (b)

Figure 4-1 – (a) SVM confusion matrix and (b) SVM ROC curve for entire field.

KNN method

This method exhibited superior accuracy relative to SVM, particularly when the model was applied for 12.25in section and whole field data. The highest evaluation measures were attained in 12.25in section, with accuracy of 89.17%, precision, recall and F1 score are 0.89. The field application demonstrated favorable outcomes with an accuracy of 81.56% and precision 0.83, recall 0.82 and F1 score 0.81. The lowest accuracy noticed in 8in section. Despite the high precision of 0.8, the levels of recall and F1 score are both modest, recorded at 0.69 and 0.65, respectively.

The ROC curve shown in Figure 4-2 (b) demonstrates the model's overall good performance, as indicated by the AUC values for each stuck type. Similar to SVM, the model correctly detects TS and OP cases with AUC 1, while NS, S and PO exhibit significantly higher performance compared to the SVM model.

The model's overall performance is satisfactory for some specific drilling intervals. However, similarly to SVM, the model may categorize certain data as noisy because of big ranges of values and being sensitive to outliers. At the same time, multi-classification problem in combination with imbalanced data remains as a significant limitation of this model.

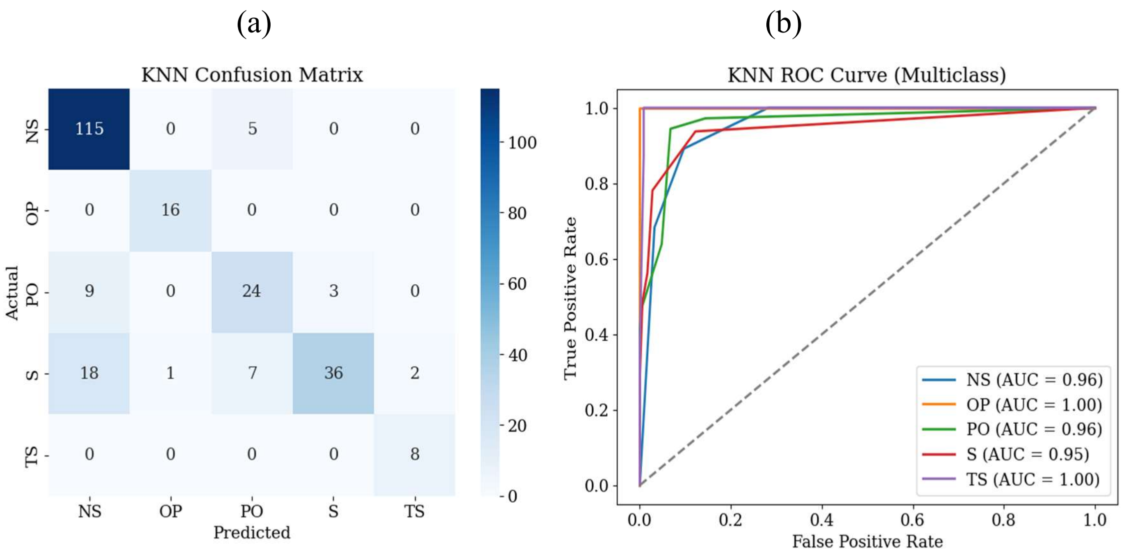


Figure 4-2 - (a) KNN confusion matrix and (b) KNN ROC curve for entire field.

LR method

Despite the simplicity of this technique, overall good performance can be concluded from the accuracy measures for each hole section. The highest results were demonstrated in 12.25in section achieving an accuracy of 95%, precision, recall and F1 score equal to 0.95. The lowest results achieved in the complete field dataset exhibited accuracy of 84.02% and other parameters of 0.84. From the ROC curve illustrated in the Figure 4-3 (b), indicates that TS, OP and PO instances were identified with the highest accuracy, achieving an AUC of 1. While S and NS has slightly lower values of AUC of 0.95, however, hey can still be regarded as outstanding results.

The model's overall performance is considered as high, with accuracy and other error measures exceeding 84%. Similar to KNN and SVM approaches, the model is influenced by noisy and imbalanced data; nevertheless, it demonstrates superior performance in handling large datasets with a high level of interpretability.

(a) (b)

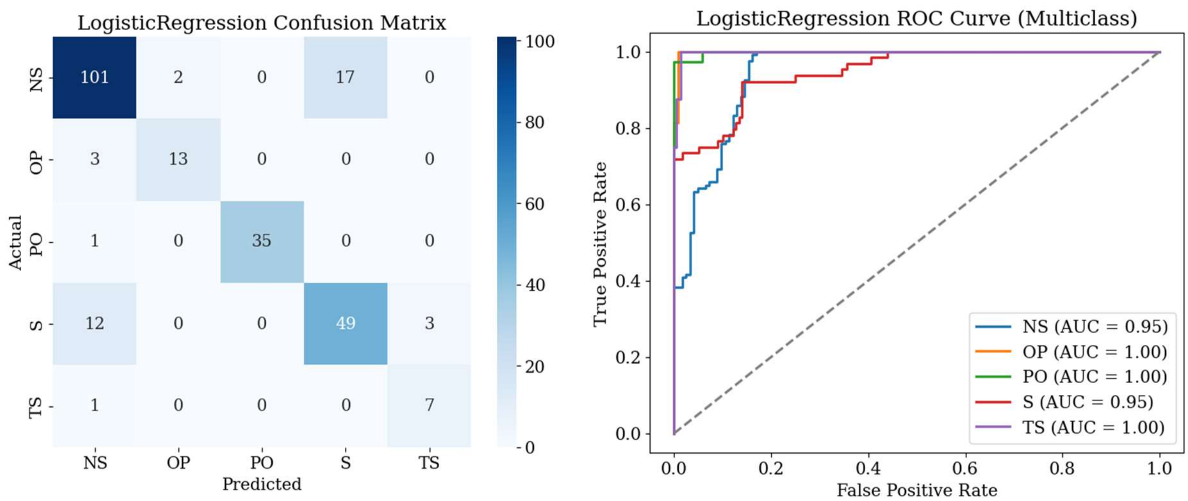


Figure 4-3 - (a) LR confusion matrix and (b) LR ROC curve for entire field.

DT method

Among all approaches used in this project DT provides one of the top performance outcomes. The accuracy of the model varies from 96.67% to 100%, while other parameters such as precision, recall and F1 score fall between 0.97 and 1 for each hole section including full data from the field. Confusion matrix of the field data illustrated in Figure 4-4(a) indicates that only one instance of NS was erroneously classified as stuck.

ROC curve shown in the Figure 4-4(b) demonstrates the model's exceptional performance, exhibiting an AUC value of 1 for each category of stuck case. The results indicate that the method has a highest applicability for stuck pipe prediction classification problem. Since the model exhibits low sensitivity to noisy data, manages imbalanced data and performs well with larger datasets.

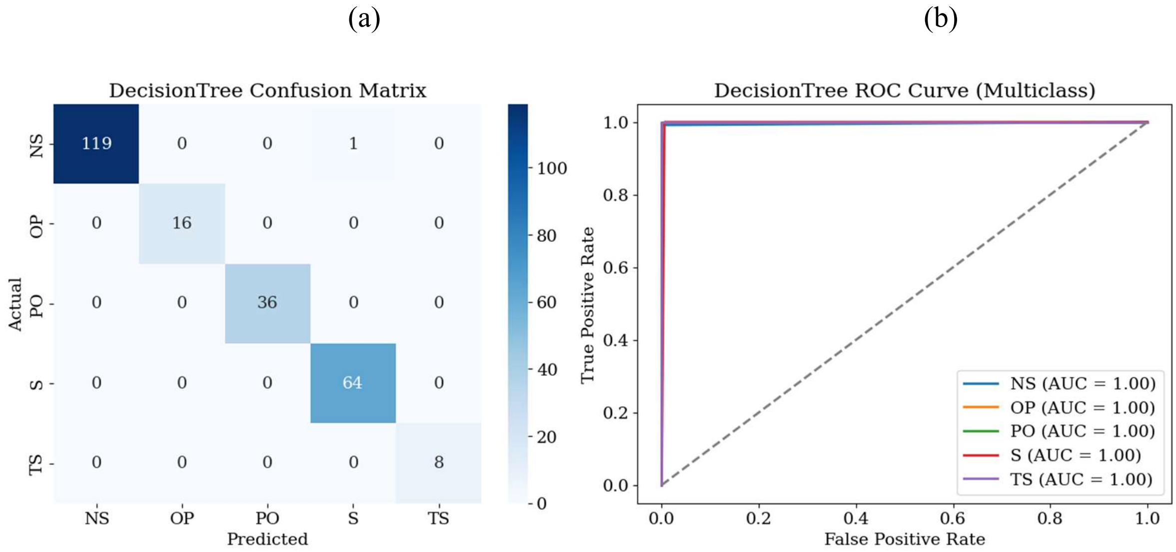


Figure 4-4 - (a) DT confusion matrix and (b) DT ROC curve for entire field.

CatBoost method

This ensemble learning method with boosting approach is regarded as one of the most accurate approaches suitable for the given case of stuck pipe prediction. It attained the utmost accuracy and error metrics for each borehole section and field data. The accuracy level ranges from 98.77% to 100%; while, other metrics vary from 0.99 to 1. Although the overall results slightly lower results in comprehensive field evaluation compared to DT approach, the individual hole sections demonstrate superior performance relative to DT. From the confusion matrix in Figure 4-5 (a) it can be observed that just 3 actual stuck cases were incorrectly predicted as non-stuck.

From the ROC curve illustrated in the Figure 4-5(b) the model achieves maximum accuracy across all stuck pipe types with AUC 1 for each case. The methodology works efficiently with outliers and noisy data and large imbalanced datasets, while delivering outstanding performance in multi-classification tasks. Overfitting of the results can be avoided with the help of internal regularization techniques, which make this methodology preferable for the given stuck pipe case.

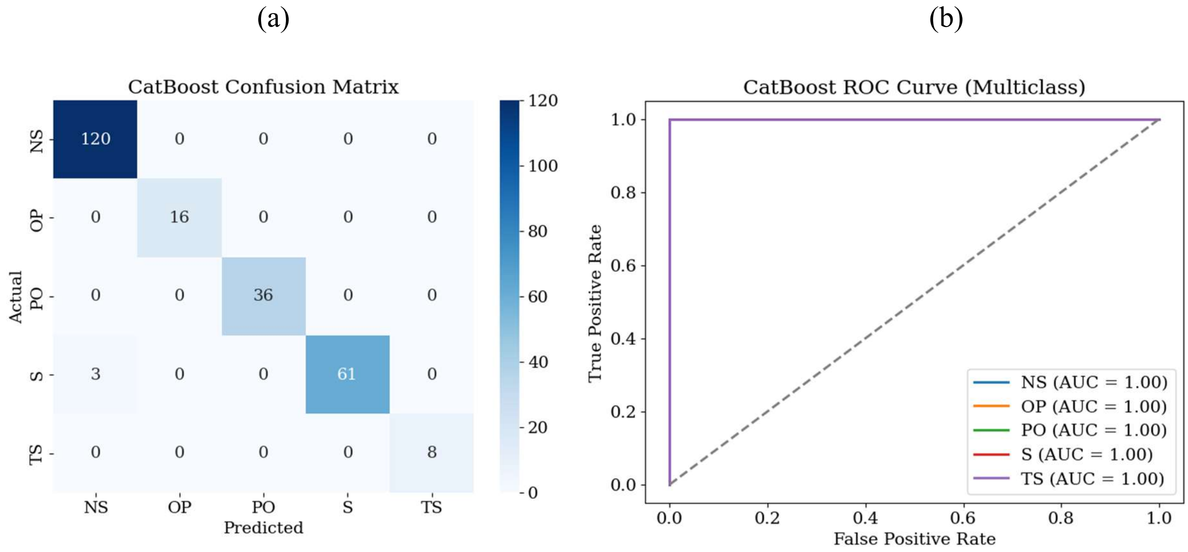


Figure 4-5- (a) CatBoost confusion matrix and (b) CatBoost ROC curve for entire field.

4.2 Model cross-validation

Five-fold cross validation method was used to evaluate the models for overfitting and assess the overall performance. Each model went through training 5 times, where data from k-1 folds (4) were used for training the models, and remaining one-fold data for validation. After each validation step, performance evaluation metrices, such as accuracy and F1 score were calculated. This step was repeated 5 times. Average error metrices were identified to assess the models' generalization ability. This validation method was applied to best performing DT-based and CatBoost-based models for each hole section separately, and for entire field data. Confusion matrix and ROC curve for the best performing folds of DT and CatBoost models are presented in the **Error! Reference source not found.**. The results of 5-fold cross validation are summarized in the Table 4-2:

Figure 4-6 – DT training and validation confusion matrixes of the best performing fold (Fold 3)

Figure 4-7 – CatBoost training and validation confusion matrixes of the best performing fold (Fold 3)

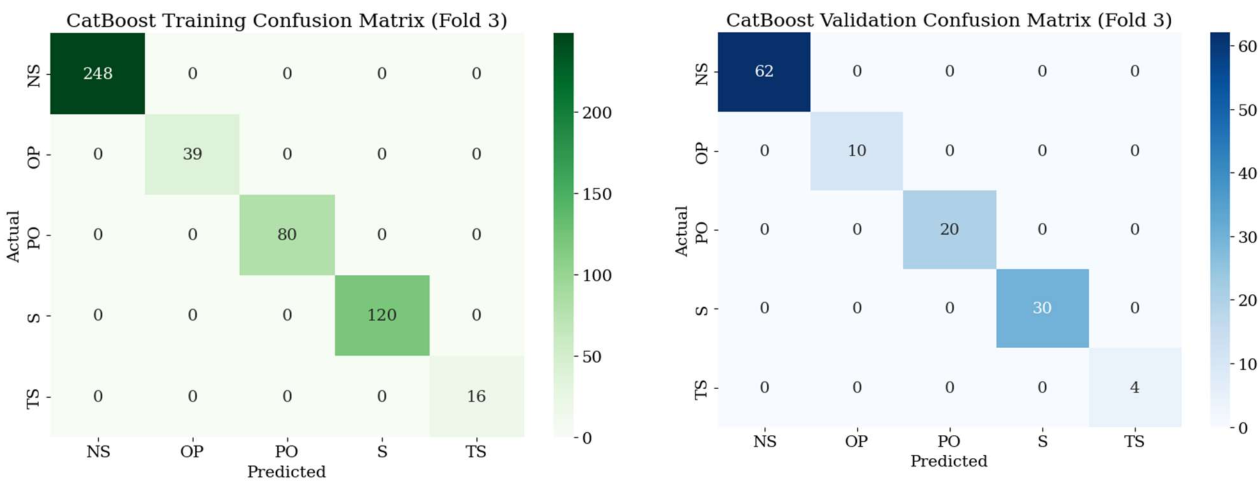


Table 4-2 – 5-fold cross validation results:

Model	Dataset	Parameter	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Average
DT	***Train***	***Accuracy***	100%	100%	100%	100%	100%	**100%**
		F1 score	1	1	1	1	1	**1**
	Test	***Accuracy***	99.21%	97.62%	100%	99.21%	99.20%	**99.05%**
		F1 score	0.99	0.98	1	0.99	0.99	**0.99**
CatBoost	***Train***	***Accuracy***	100%	100%	100%	100%	100%	**100%**
		F1 score	1	1	1	1	1	**1**
	Test	***Accuracy***	100%	99.21%	100%	100%	99.20%	**99.68%**
		F1 score	1	0.99	1	1	0.99	**1**

From the results of cross validation, it can be seen that accuracy and F1 score values for both training and testing phases of DT and CatBoost models in each fold are close to each other, and no significant discrepancy observed in the calculations. The same results observed in the values between the folds, which indicates that the models are not demonstrating any signs of overfitting. In case of overfitting, it would be expected to see notable difference between the training and testing results or in the calculations among the folds.

4.3 Model performance comparison

Comparison of obtained intelligent models with existing models from recently published studies was conducted using error and performance evaluation metrices to assess the overall modeling performance of developed smart models and compare key parameters derived from current study. This analysis is conducted based on three main parameters: accuracy, F1 score and AUC. Performance metrics show variations based on ML algorithm selection which creates challenges for model performance assessment. Additionally, quantity of datasets, input parameters, data sources significantly vary in each case. Considering the differences of these methodologies, a comparative analysis of key metrics can be done in order to evaluate current ML models. As shown in Figure 4-8, the overall performance of the proposed ML-based models outperforms pure ANN models. It was observed, that ANN models demonstrate high accuracy, however their F1 score and AUC values are lower compared to ML methods. F1 score is an important parameter that combines precision, which demonstrates positive predictions accuracy, and recall, which counts for actual positive outcomes, providing a more comprehensive assessment of existing models. Current project attained the highest F1 score in DT and CatBoost models, notably higher compared to several other models. AUC derived from ROC is another critical parameter indicating the model's ability to differentiate between positive and negative cases. This research attained the highest AUC score for DT and CatBoost models in comparison with other studies. It can be also highlighted, that LR, KNN and SVM models exhibit lower performance relative to other studies, despite their enhanced ability to distinguish between positive and false cases, they lack of accuracy.

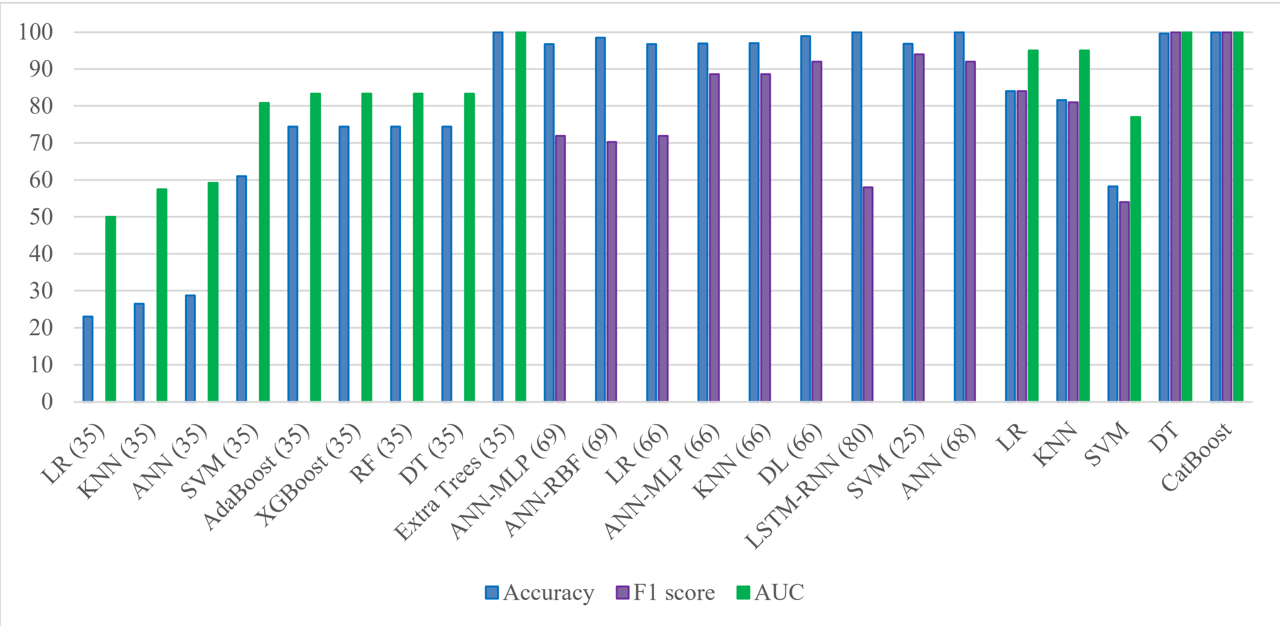


Figure 4-8 – Performance comparison graph of current project models with existing studies from literature review.

4.4 Interpretability of the models

Ability to interpret outputs of intelligent models is a critical aspect for understanding the insights of the model, transparency of evaluation and level of trust for obtained results. Despite the understanding of algorithms behind ML approaches, the models used these days remain as a black-box for their users (Antonini et al., 2024). Hence, it is important for the chosen intelligent model to achieve high level of output accuracy, at the same time to demonstrate interpretability of results. Number of techniques are used these days for models' outcomes interpretation, which mainly can be divided into global interpretation techniques and local explanation methods (Lundberg et al., 2019).

Global interpretation techniques usually focus on defining the features that has the most significant effect on general prediction based on entire model. While local explanation methods focus on each individual prediction to understand what is the impact of input data for each sample output. This method is not widely used compared to global interpretation because of its complexity, especially in complex classifiers. However, local interpretability gives a valuable insight to each model and allow the user to understand why this specific decision was made by intelligent model (Antonini et al., 2024).

To further understand the overall behavior of main classifier, and define the most influential parameters in the contest of individual hole sections and entire field, global interpretation technique known as feature importance analysis with using Loss Function Change method was used for this project. At the same time, local explanations have been also used to analyze each decision that was made by the model and how input parameters affect them. This analysis was performed with the use of SHAP (SHapley Additive exPlanations) methodology to evaluate individual predictions on each output.

4.5 Feature importance evaluation

Feature importance analysis is one of the essential steps of modeling evaluation process. It helps to define the most influential parameters of the model which allow to access valuable insights of the stuck pipe condition. By understanding them we can enhance the model performance and increase accuracy of results, at the same time it allows to identify possible root cause of each stuck situation and define the correct strategy to eliminate these risks in the future field development stages.

The feature importance analysis was performed with the use of Loss Function Change method as a global interpretation of intelligent model. This method in the modeling determines feature importance through analyzing how each parameter affects the training loss reduction process. This methodology computes an approximation of how loss values differ between an original model and a modified variant which does not contain a specific feature. During exclusion it removes the feature from all trees so the model does not require a full retraining process. Positive value shows that a given feature plays a critical role in reducing model loss resulting in better performance. While a negative score indicates that activating the feature may cause the overall growth of the loss, which implies detrimental performance from that feature. The main reason for that output could be related to underfitting issues or significant noise presence in the input data.

The feature evaluation has been performed for CatBoost model, since this model was evaluated as the most powerful based on accuracy and error estimations, and concluded as one of the most suitable models for given stuck pipe prediction problem. This analysis performed to define the feature importance on different type of stuck condition, including S – stuck, NS – non-stuck, PO – pack off and OP – overpull. And applied to evaluate different hole section which exhibited the most cases of stuck condition – 16in and 12.25in sections, to define the most contributing parameter and understand the root cause of each condition. 16in and 8in sections were eliminated from the individual evaluation because of limited number of stuck cases, since only 1 and 2 stuck issues detected respectively, which does not allow for comprehensive analysis and the results may lead to biased or incorrect conclusions. However, it was decided to include the outputs of these sections into entire field evaluation part. The final evaluation was made on the whole field data.

16 inches hole section

This hole section is considered as one of the problematic hole sections which exhibited 14 different stuck situations. The aim of this section is to isolate top layered sedimentary formation prior to entering massive salt intervals. The interval is characterized by its layered structure of sandstone and argillite, which is known as a fine-grained rock composed of clay particles (Tinseau et al., 2006). The feature importance evaluation of the model output for each stuck and non-stuck condition is presented in the form of bar plots in Figure 4-9 and (b)

Figure 4-10. Where on X axis is feature importance parameter calculated based on Loss Function Change method represents the contribution of each parameter to the output of the model. Input features are presented on Y axis with the highest contribution listed from top to bottom. A summary of top five the most contributing parameters to the classifier outputs depending on stuck condition is presented in Table 4-3.

Table 4-3 – Top contributing parameters for different stuck classes based on feature importance evaluation of the model (16 inches section):

Number of parameters	***Non-stuck***	***Stuck***	***Overpull***	***Pack-off***
1	Sandstone content	Flow rate (MFIN)	Depth (DPTS)	Plastic Viscosity (PV)
2	Flow rate (MFIN)	Salt content	Clay content	Flow rate (MFIN)
3	Depth (DPTS)	Stand Pipe Pressure (PSPM)	Mud weight (MW)	Clay content

4	Clay content	BHA length (BL)	Solid content	Sandstone content
5	RPM	Gel Strength (GEL10MIN)	Stand Pipe Pressure (PSPM)	Mud weight (MW)

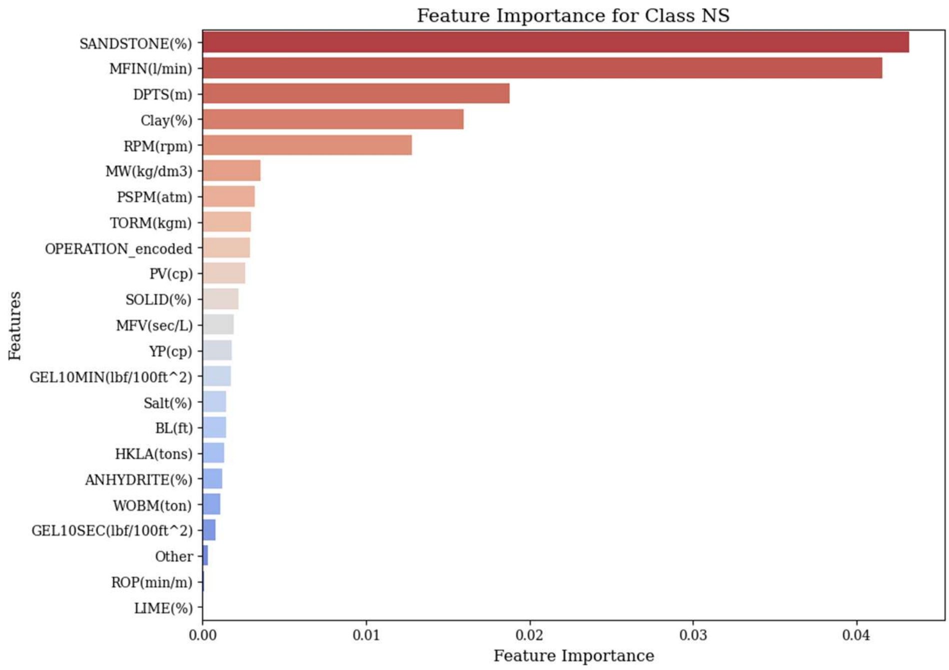
Sandstone content has the highest feature importance value among others on the non-stuck condition. Based on that and analysis of geological and drilling data, it was concluded that non-stuck normal activity is mainly took place when the formation was predominantly sandstone with no or small inclusion of clay rocks. As the clay rocks are known for their ability to swell when in the contact with water causing the tightness of borehole that may lead to stuck pipe situation (Murtaza et al., 2020). Similarly, clay content considered as one of the highest contributors to pack-off and overpull conditions ((b)

Figure 4-10 (a) and (b)).

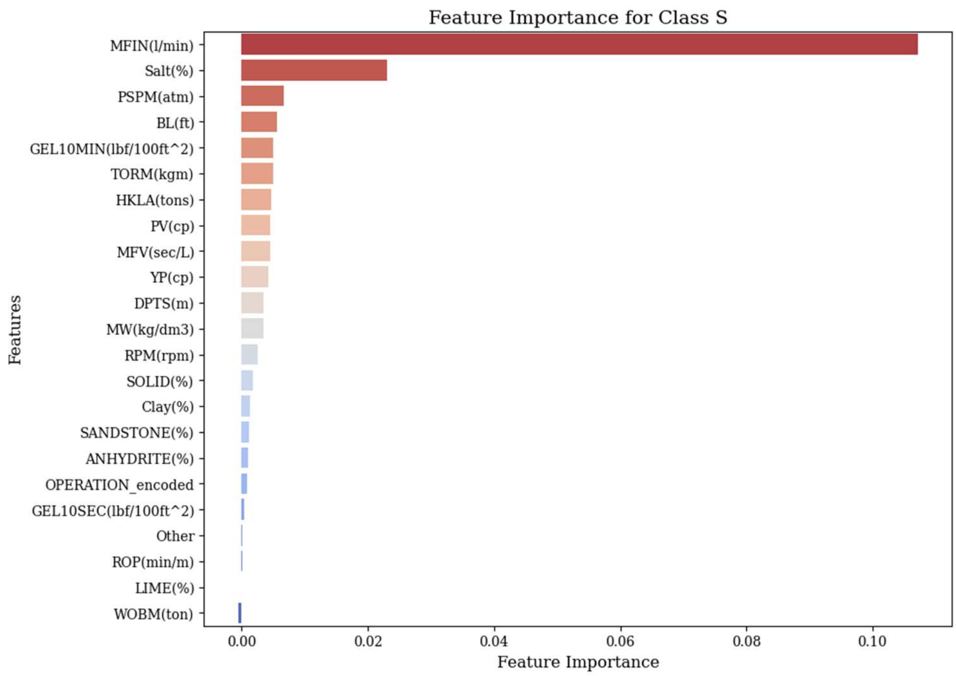
Flow rate has also high value of feature importance on non-stuck, stuck and pack-off situations. By analyzing drilling data, it was noticed that this section has the highest number of pack-off induced stuck pipe conditions among other sections. Meaning that the most of the stuck situations was characterized not only with inability to move the string, but also with restricted flow between drill string and borehole, which is resulted in clay swelling effect that blocked the annular space. Hence, ability to pump the fluid has been picked by the model as an important feature contributing to non-stuck and pack-off conditions. It also can be noted, that depth has been considered as an important parameter affecting non-stuck and overpull conditions. From the analysis, it was depicted that the non-stuck conditions mainly characterized by shallower depths compared to overpull conditions.

Some presence of salt rocks also observed while drilling this section, which affected the stuck pipe issues. As it seen that the salt is a second contributing factor to the stuck situation (Figure 4-9(b)).

Interesting observation is made in the feature importance analysis of pack-off condition, that plastic viscosity (PV) has the highest effect on the model output. This parameter is known as a significant feature demonstrating efficiency of drilling mud to transport the cuttings to the surface and preventing accumulation of cuttings in the borehole (Al-Malki et al., 2023). The high value of this parameter demonstrates that poor hole cleaning is leading to pack off the drill string. Similarly, mud weight and solid content was picked by the model as one of the primary influential parameters on overpull events, as these rheological properties of the mud demonstrate efficiency of the hole cleaning process.

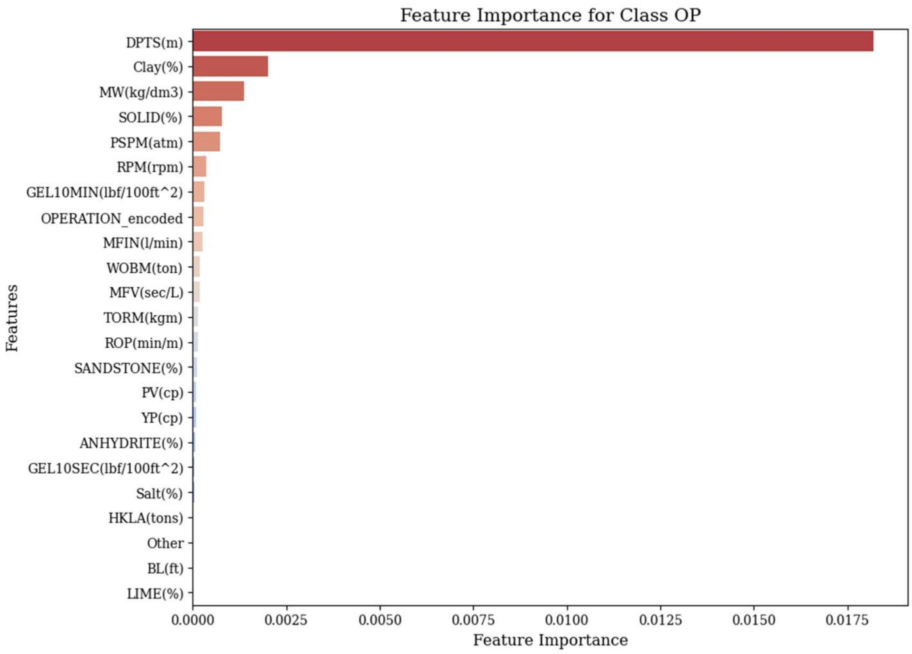


(a)

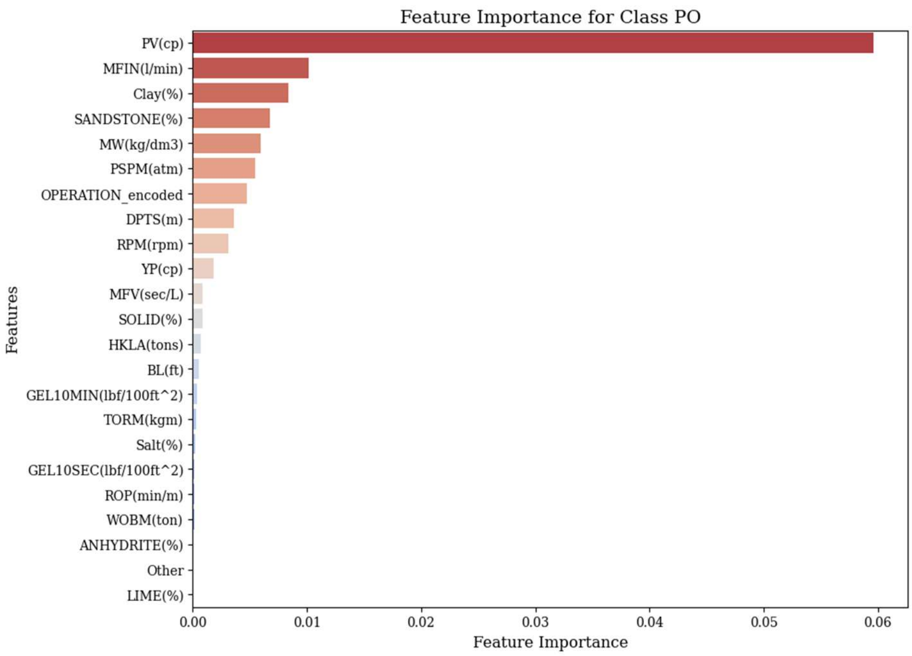


(b)

Figure 4-9 – 16 inches section feature importance evaluation plots: (a) non-stuck (NS), (b) stuck (S).



(a)



(b)

Figure 4-10 – 16 inches section feature importance evaluation plots: (a) overpull (OP), (b) pack-off (PO).

12 inches hole section

This hole section is the most challenging sections of the well, as the main goal of this section is to drill through massive salt intervals and isolate it by running intermediate casing. As previously highlighted, because of its ability to creep a drilling through salt intervals represent notable challenges to drilling process imposing number of problems, including stuck pipe condition. The model's output was evaluated for the feature importance in each case of stuck and non-stuck conditions and illustrated as bar plots in

(b)

Figure 4-11 and (b)

Figure 4-12. the most effecting parameters to the model results based on stuck situations are summarized in Table 4-4.

Table 4-4 – Top contributing parameters for different stuck classes based on feature importance evaluation of the model (12 inches section):

Number of parameters	*Non-stuck*	*Stuck*	*Overpull*	*Pack-off*
1	RPM	Salt content	Clay content	Gel strength (10sec)
2	Anhydrite content	RPM	Depth (DPTS)	Yield Point (YP)
3	Salt content	Anhydrite content	Yield Point (YP)	Stand Pipe Pressure (PSPM)
4	Flow rate (MFIN)	Flow rate (MFIN)	Anhydrite content	Plastic Viscosity (PV)
5	Depth (DPTS)	Clay content	Salt content	Marsh Funnel Viscosity (MFV)

As it was expected salt content acts as a main contributor to the outputs of the model for stuck and non-stuck conditions ((b)

Figure 4-11 (a) and (b)). Previously it was discussed, that nature of evaporite formations and their ability to flow, especially under hard downhole condition, present a significant risk of pipe stuck issue. Problems with stuck pipe situation in this section is directly related to formation type, as we can see from the plots in

(b)

Figure 4-11 and (b)

Figure 4-12, anhydrites also significantly affect the model results. Additionally, presence of clay minerals is the most contributing factor to the overpull problems encountered in this interval ((b)

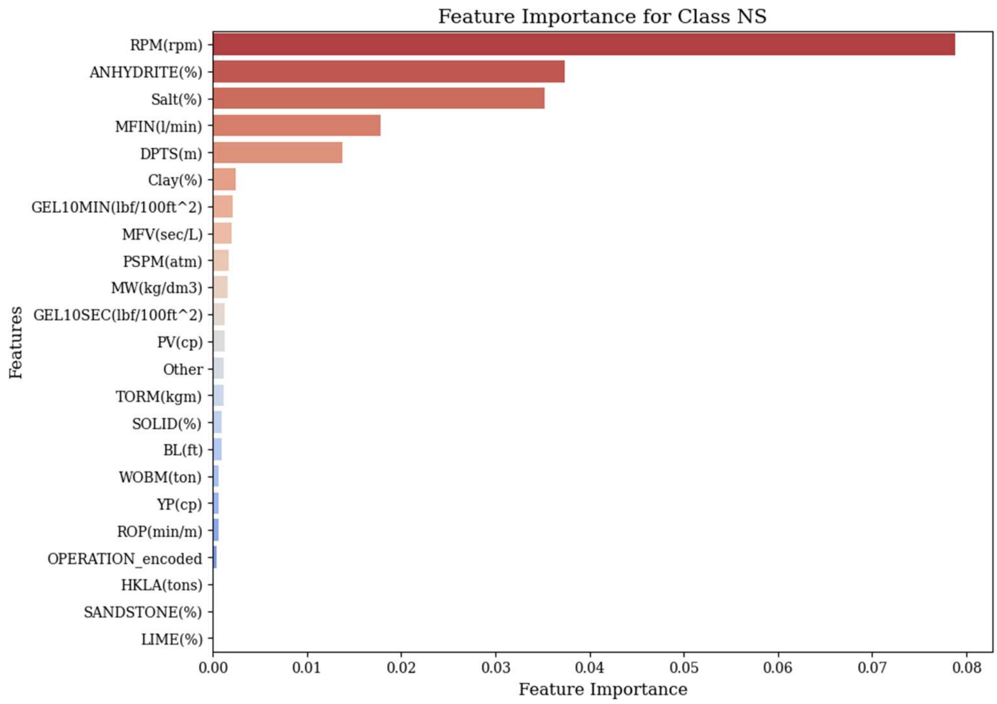
Figure 4-11(a)).

From the drilling parameters side, it is notable to mention the effect of RPMs on the output of non-stuck and stuck modeling conditions. It is mainly related to the fact, that RPM is important drilling parameter that reacts to stuck condition, as we can see that the drill string is limited in rotational motion or its exhibiting difficulties with rotation showing fluctuations of readings, similarly to surface torque behavior.

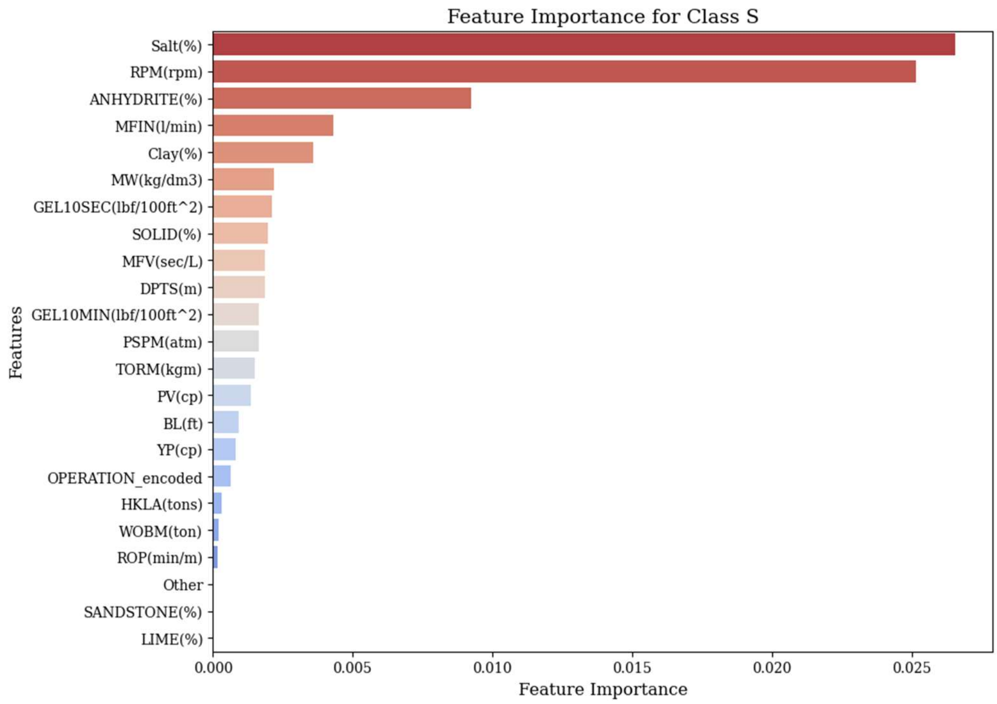
From the feature importance graph of overpull class (
(b)
Figure 4-12(a)), similarly to 16 inches hole section, main contributor was defined as clay content because of its reactivity to water and swelling effect. In addition, depth value, analogously to 16 inches section, acts as one of the primary influencers to overpulls. From drilling data, it was depicted that the risk of overpull increases notably with depth.

Rheological parameters of drilling mud, such as gel strength, yield point, plastic and funnel viscosities, were identified as the main drivers for model results representing pack-off class ((b)
Figure 4-12(b)). As it was depicted previously, these parameters illustrate effectiveness of mud system to maintain the hole stability and cuttings removal, which in this case is not sufficient, as it causing pack off situations.

Interesting observation can be made that in this section the flow rate is not the most contributing parameter compared to previous interval, which can be because of the fact that the formation changed to predominantly salt and clay rocks and not affected by flow rate unlike 16 inches hole.

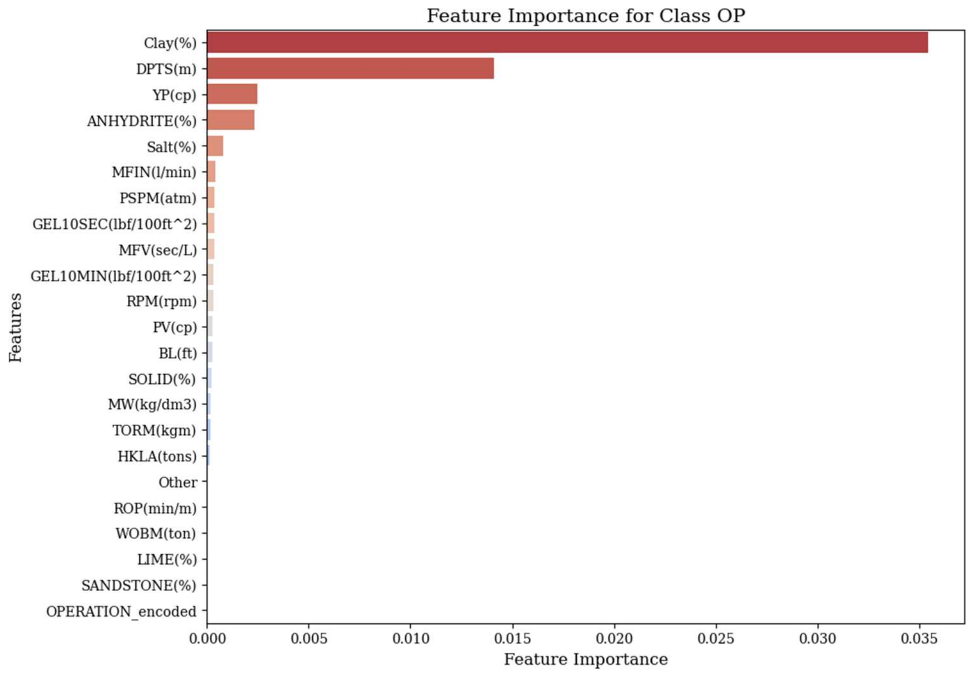


(a)

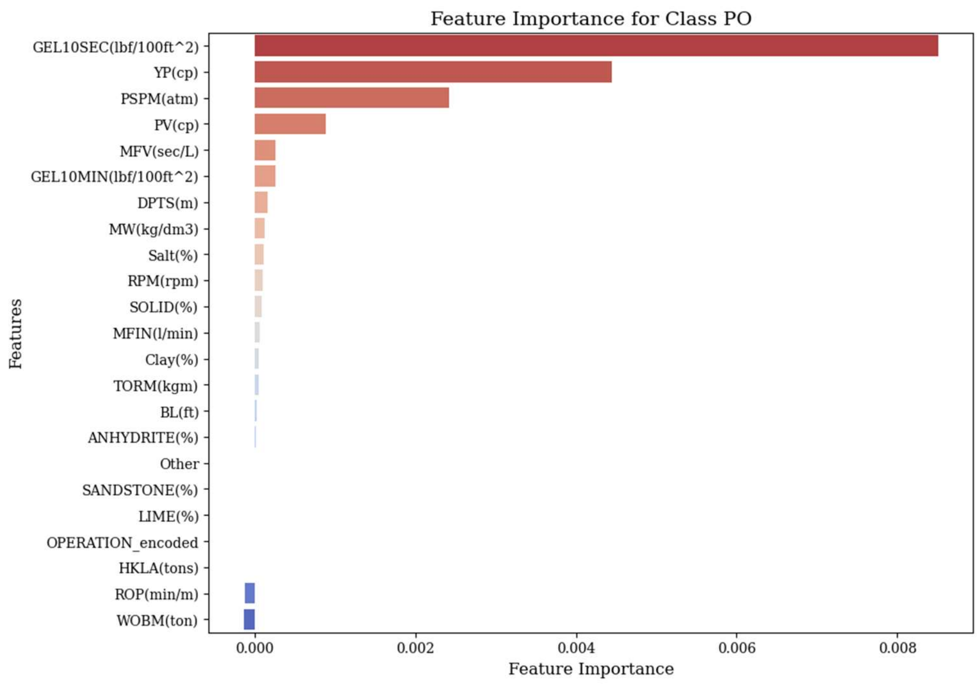


(b)

Figure 4-11 – 12 inches section feature importance evaluation plots: (a) non-stuck (NS), (b) stuck (S).



(a)



(b)

Figure 4-12 – 12 inches section feature importance evaluation plots: (a) overpull (OP), (b) pack-off (PO).

Entire Field

Feature importance evaluation of the modeling based on entire field data gives an opportunity to see the whole picture of stuck pipe problem of this particular field. To understand, the main parameters affecting the modeling when the model is applied to entire dataset. And define the most and the least important parameters contributing to stuck problem. Significance of this evaluation is also depended on the presence of top section of 26in and production section 8in. The field evaluation of the model performance based on feature importance analysis is illustrated in (b)

Figure 4-13 and Figure 4-14. In addition, the primary parameters that affect the modeling outcomes based on stuck classification are summarized in Table 4-5 – Top contributing parameters for different stuck classes based on feature importance evaluation of the model (entire field):.

Table 4-5 – Top contributing parameters for different stuck classes based on feature importance evaluation of the model (entire field):

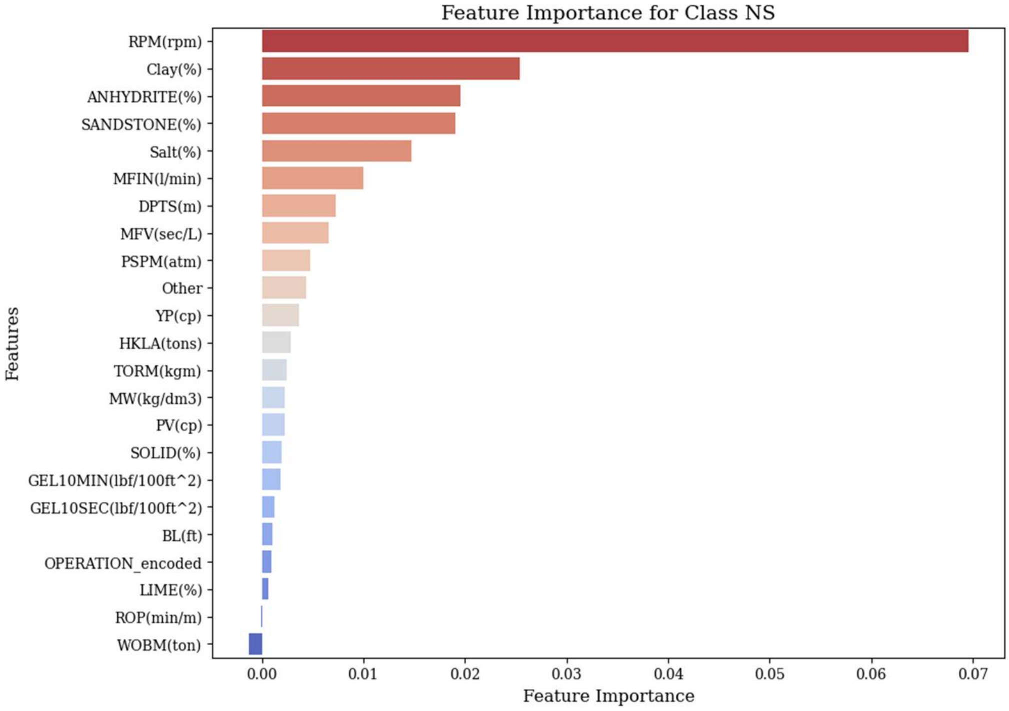
Number of parameters	***Non-stuck***	***Stuck***	***Overpull***	***Pack-off***	***Tight Spot***
1	RPM	Salt content	Clay content	Yield Point (YP)	Marsh Funnel Viscosity (MFV)
2	Clay content	RPM	Operations	Clay content	Stand Pipe Pressure (PSPM)
3	Anhydrite content	Flow rate (MFIN)	Depth (DPTS)	Plastic Viscosity (PV)	Flow rate (MFIN)
4	Sandstone content	Marsh Funnel Viscosity (MFV)	Stand Pipe Pressure (PSPM)	Depth (DPTS)	Plastic Viscosity (PV)
5	Salt content	Gel strength (10sec)	Solid content	Flow rate (MFIN)	Operations

This evaluation allows to conclude that the stuck situations at the field level is clearly formation related problem, as it can be depicted from all the plots of (b)

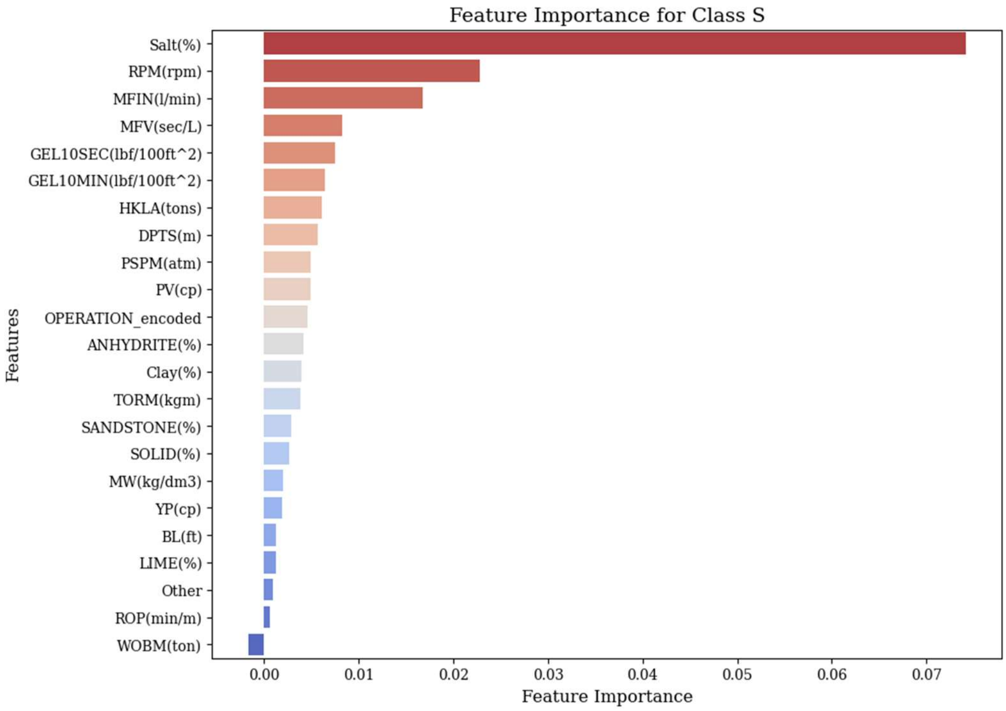
Figure 4-13 and Figure 4-14 that salt and clay formations are one of the main contributors to stuck pipe condition. Contribution of salt content to stuck classification is several times higher than any other parameter, despite the fact, the salts are mainly presented in intermediate 12.25 inches section. Problems of overpull of this field mainly rely on presence of clay minerals in the formation drilled (Figure 4-14(a)). As it is notably higher than any other parameter. Pack-offs on this field are primarily resulted in significant presence of clay in the formation and its ability to swell under downhole conditions (Figure 4-14(b)).

From the surface drilling parameters, it can be concluded that RPM and flow rate are the main factors affecting the model output on classifying stuck and non-stuck cases. As these two parameters show the ability of pipe to rotate and pump the fluid down the hole. Hence, they can be considered as main parameters for monitoring and detecting stuck condition.

It can be noted, that rheological properties of the mud are critical in data modeling of stuck conditions, especially in classification based on stuck mechanisms. Therefore, it is seen that yield point and plastic viscosity are top features influencing the model outcomes. Additional evaluation of tight spot class is showing the importance of mud viscosity measurement (MFV) combined with drilling parameters of flow rate and stand pipe pressure can indicate that mud system is critical in tight spot elimination while drilling (Figure 4-14(c)).

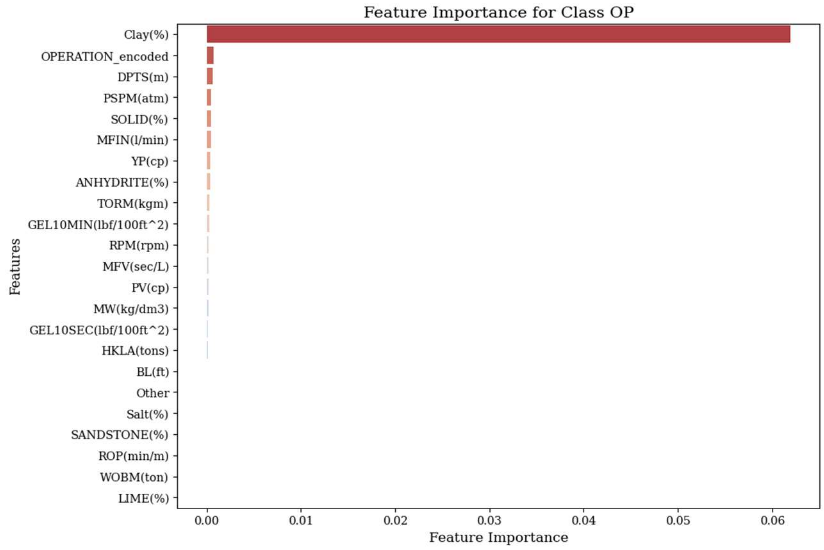


(a)

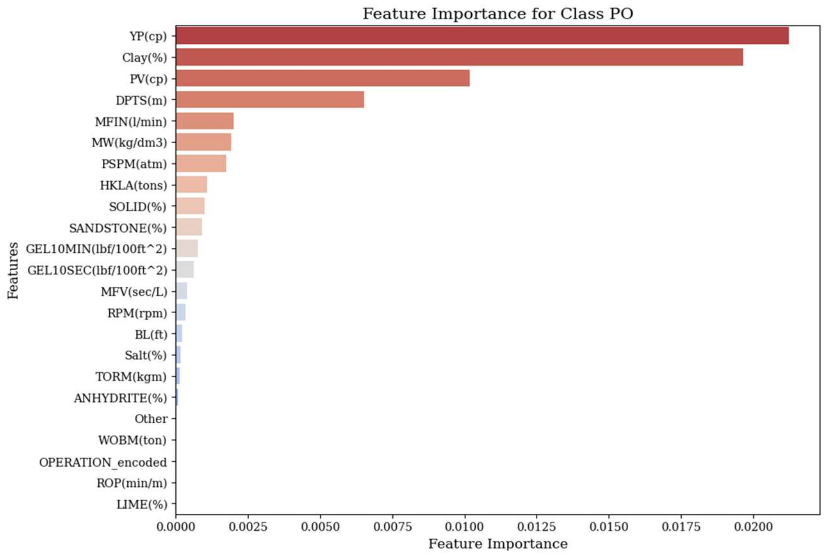


(b)

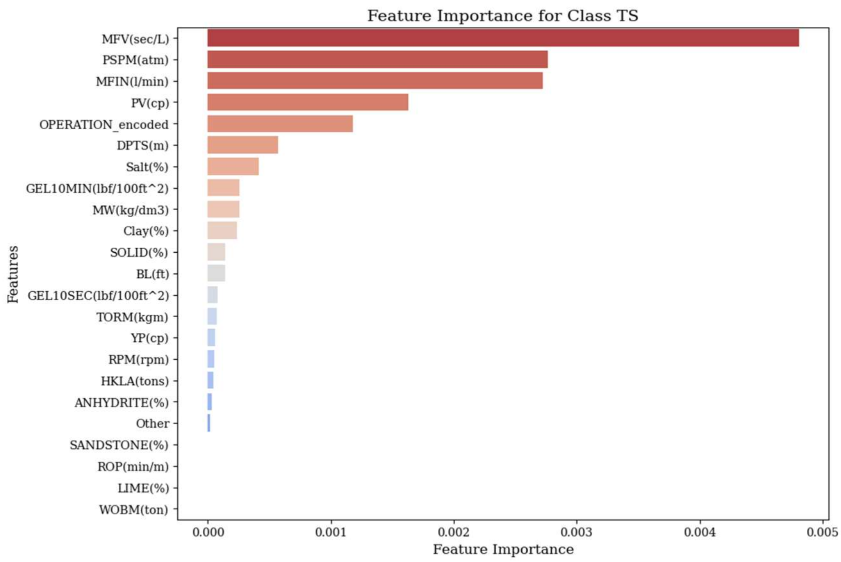
Figure 4-13 – Entire field feature importance evaluation plots: (a) non-stuck (NS), (b) stuck (S).



(a)



(b)



(c)

Figure 4-14 – Entire field feature importance evaluation plots: (a) overpull (OP), (b) pack-off (PO), (c) tight spot (TS).

4.6 Local explanation using SHAP plot

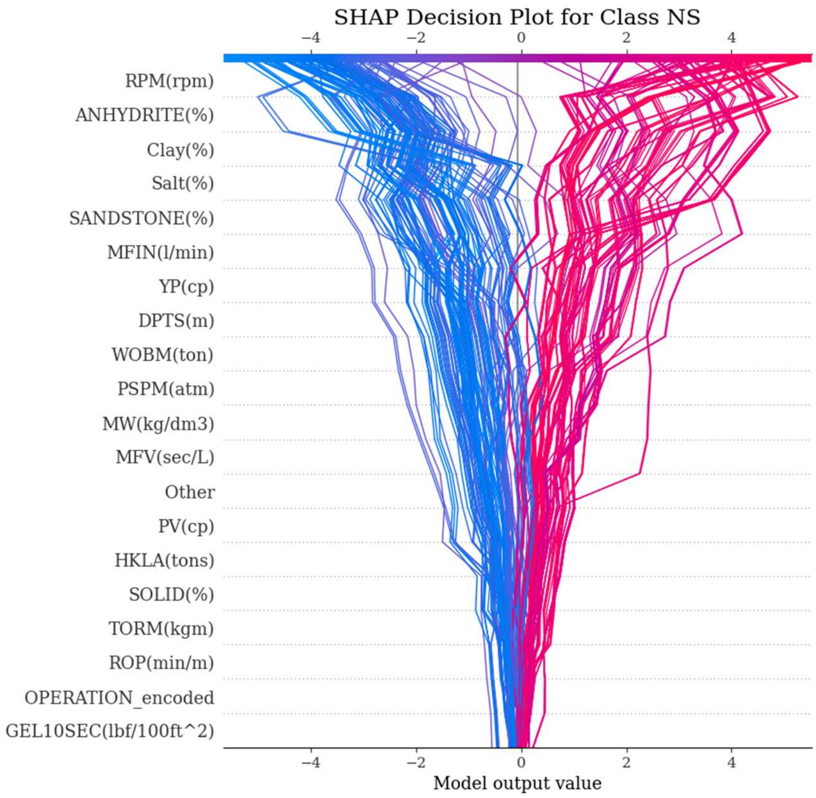
The further evaluation of modeling results was continued with local explanation analysis using SHAP methodology. This technique is based on Shapley values obtained from cooperative game theory developed by Shapley in 1951. The main concept of the method defines a fair distribution of credits among the group of players in cooperation depending on their contribution. In the case of ML, SHAP analysis allows to assess individual contribution of each parameter and their effect on the final output with the help of calculated Shapley values with local explanations on each sample output. SHAP plots allow to visually evaluate SHAP values based on individual contribution of each parameter and perform a local explanation of each individual output sample (Li, 2022).

This evaluation was performed for each model's output of individual hole sections and entire field. SHAP decision plot lines show how individual predictions are affected by different features while the model chooses a certain class based on classifier output. As seen from Figure 4-15, Figure 4-16 and

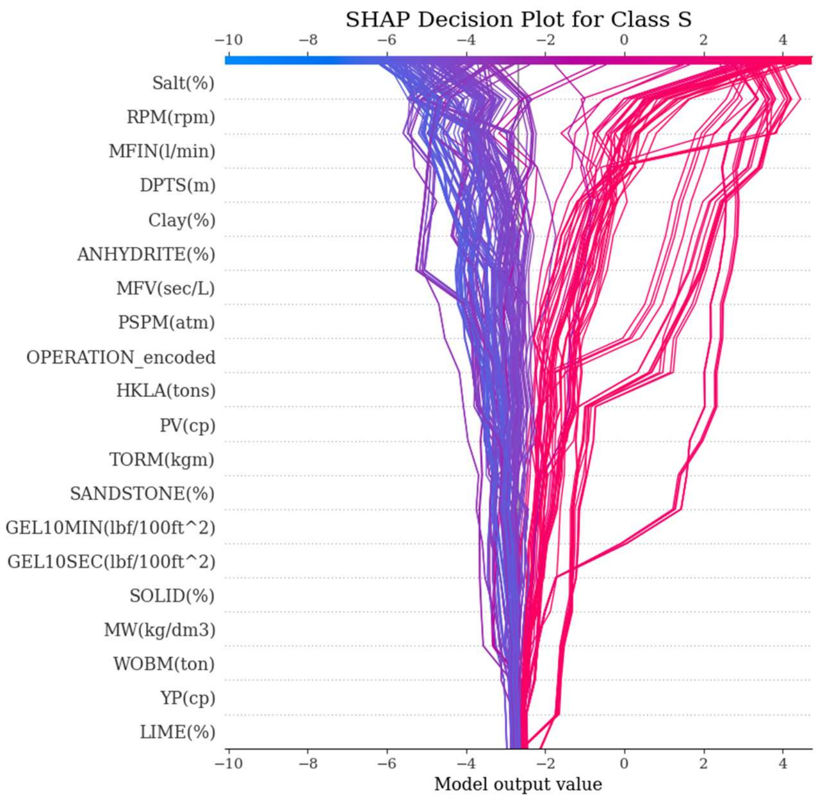
Figure 4-17 the SHAP plots were developed for each class output of entire field data. The SHAP value of each instance accumulates along each line to demonstrate how feature changes affect the likelihood of classifying a data point into a certain class. Each horizontal line in the plot represents a feature which shows the effect of specific instance values on changing the model prediction deviation from base value (average model output across dataset). The SHAP values explained through the x-axis show the strength along with direction of influence that each feature generates. The values right of 0-axis vertical dashed line (positive SHAP) indicate features that advance the prediction toward the certain class; however, negative values (left of 0) mean features that lower this specific class probability. SHAP contributions of each feature rise along each line until the upper position is reached. For instance, the features located highest on the list starting with RPM and including Anhydrite and Clay contents together with

Salt and Sandstone play the biggest role in determining the classification outcome as Non-Stuck (Figure 4-15(a)). Lines crossing in the plot indicates that various cases have distinct reactions to a specific feature. Nonlinear relationships exist between features because other variables influence their effect. Each data point travels through the model decision system as a distinctive line which illustrates which features help and which ones harm the classifier when determining a class NS. A wide distribution of lines around specific features demonstrates significance as well as variability of these features when reaching the final outcome.

Similarly, decision plot illustrated in Figure 4-15(b) shows the feature effects that lead to class Stuck (class S) predictions from the model. The length and placement of horizontal lines represent features in the plot where x-axis (-10 to +4) values determine both magnitude and direction of prediction influence. Features that boost the probability of class S explain positive SHAP values present to the right of 0, while those lowering stuck likelihood display negative SHAP values located to the left of 0. The area between -10 to +4 in the plot indicates that different features have high negative strength in determining class S. The plot indicates that class S predictions rely mainly on salt content alongside operational metrics, such as RPM and flow rate, and also consider geological depth and clay composition as additional secondary indicators. The wide negative range of the x-axis shows that some features which do not appear on the top 5 list actively oppose class S predictions yet the number of positive parameters prevail. Salt content together with mechanical operational variables form key decision factors in class S classification through the model's logic.



(a)



(b)

Figure 4-15 – SHAP decision plots for each stuck pipe class evaluation based on whole field data: (a) non-stuck (NS), (b) stuck (S)

As once can see from Figure 4-16, operational metrics such as flow rate and pump pressure together with fluid rheology measurements of viscosity become the main determinants for Tight Spot classification predictions, while geological variables like depth and salt content offer supplementary support.

Predictions for class Overpull (class OP) presented in

Figure 4-17(a) are mainly influenced by clay content, and drilling parameters, such as flow rate and pump pressure, depth and fluid rheological characteristics act as supplementary signals. The SHAP value distribution reveals an asymmetry pattern, which proves that the model makes reliable class OP predictions through the use of these five essential features. As shown in

Figure 4-17(b), fluid rheology through PV and YP together with geological composition of clay and depth measurements act as primary predictors for class PO alongside operational input of flow rate. Other parameters cannot compensate for class PO because the top features have the strongest influence on model predictions according to the SHAP value distribution.

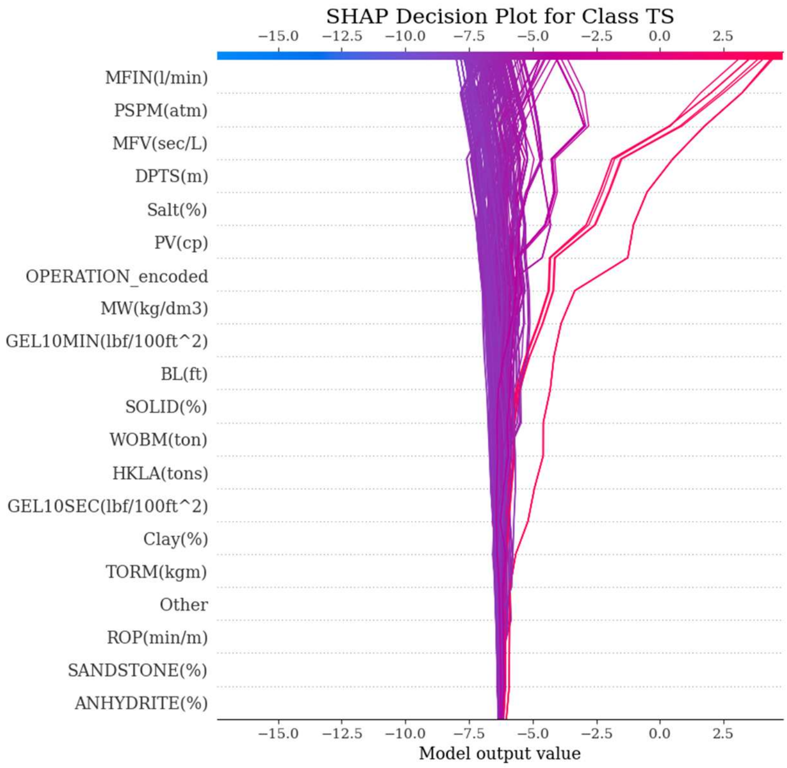
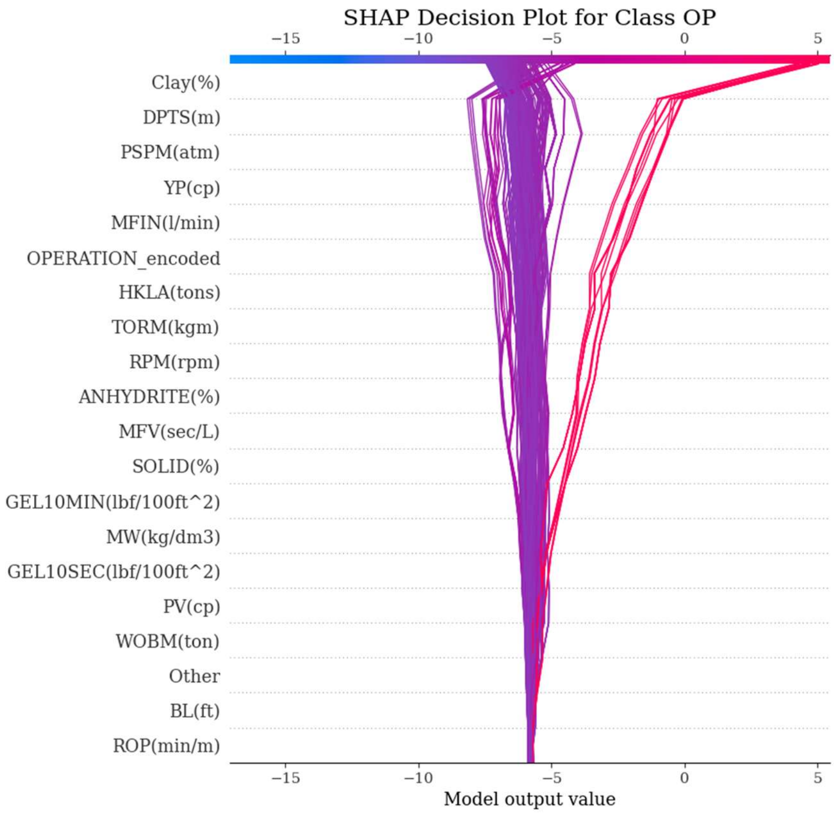
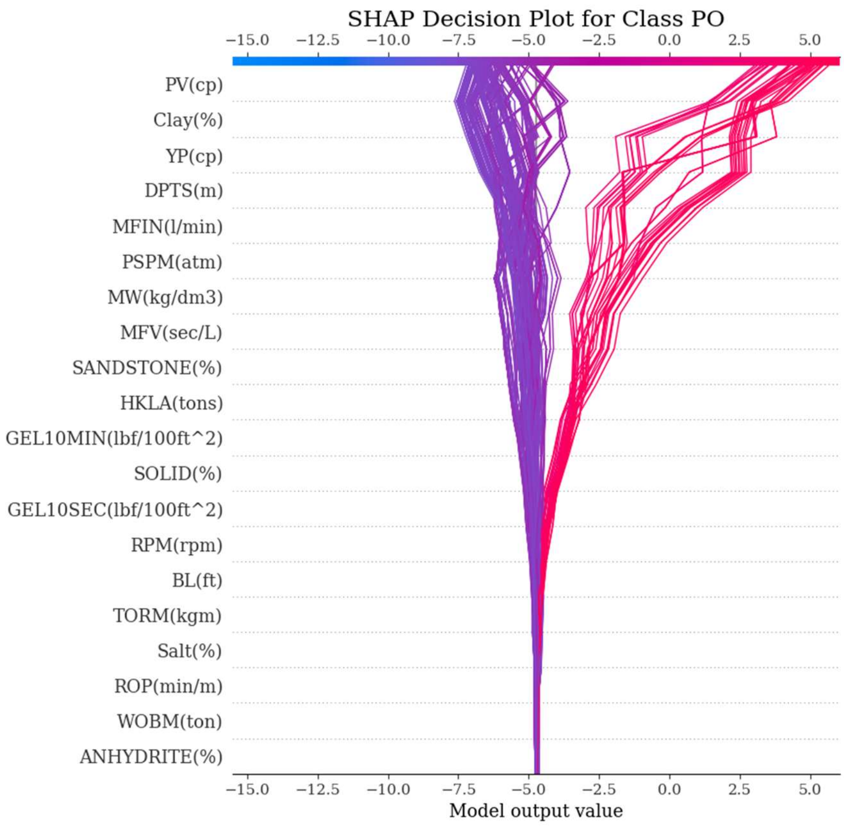


Figure 4-16 – SHAP decision plots for each stuck pipe class evaluation based on whole field data – tight spot (TS)



(a)



(b)

Figure 4-17 – SHAP decision plots for each stuck pipe class evaluation based on whole field data: (a) overpull (OP), (b) pack off (PO)

From comparison of global interpretation approach with local explanation method executed by SHAP analysis, the following similarities and differences were reviled (Table 4-6):

Table 4-6 – Comparison table of global evaluation using feature importance analysis with local explanation based on SHAP evaluation for entire field data:

Class	***Similarities***	***Differences***
Non-Stuck (NS)	- RPM - Clay content - Sandstone content	SHAP: - Anhydrite content - Salt content Feature Importance: - Flow rate - Yield Point
Stuck (S)	- Salt content - RPM - Flow rate	SHAP: - Clay content - Depth Feature Importance: - Mud weight - Gel strength (10 min)
Overpull (OP)	- Clay content - Depth - Pump Pressure	SHAP: - Flow Rate Feature Importance: - Operations - Solid content
Pack-off (PO)	- Plastic Viscosity - Clay content - Yield Point - Flow rate	SHAP: - Depth Feature Importance: - Mud Weight
Tight Spot (TS)	- Flow Rate - Pump Pressure - Depth	SHAP: - Salt content Feature Importance: - Plastic Viscosity

The following summary and conclusions can be drawn from comparison of two evaluation techniques and their impact on overall model performance and key parameters identification for drilling operations efficiency in terms of stuck pipe condition:

- The measurements of RPM and clay content demonstrate significant impact on drilling conditions throughout both categories of analysis.
- The SHAP analysis evaluates the impact of each variable at specific cases, while feature importance analysis provides global observation throughout the dataset.
- The classifications of stuck pipe issues often include Depth and Flow Rate measurements as main contributors.
- The ranking of rheological fluid properties, such as PV, YP and MW, is higher according to feature importance analysis, whereas SHAP assigns higher priority to depth-related and flow-related variables.
- SHAP analyzes Salt content as the primary influencer, even though it does not consistently appear in the feature importance evaluations of individual hole sections, which indicates local dominance of formation effects. That certainly related to presence of salts while drilling 12.25in section and its significance in stuck pipe occurrences.

5. CONCLUSIONS AND RECOMMENDATIONS

In this research work, stuck pipe condition prediction while drilling through deep and complex evaporite formations was addressed using ML-based data analytics approaches for early stuck pipe evaluation and risk interpretation purposes. The research methodology involved comprehensive field data acquisition followed by data pre-processing steps for quality control by detecting missing data and outliers, and distribution of important drilling parameters were analyzed. A predictive evaluation of ML algorithms required testing different ML algorithms, such as Logistic Regression, K-Nearest Neighbors, Support Vector Machines, Decision Trees and CatBoost through training procedures. Multiple performance metrics were used to assess the models to guarantee their reliability and robustness, such as accuracy of the outputs, precision, recall and F1 score obtained from confusion matrix. Additionally, ROC evaluation technique was used to address the limitations of each model. Feature importance assessment demonstrated how drilling parameters with geological data, and drilling fluid characteristics significantly affect stuck pipe occurrences. At the same time, SHAP analysis brought additional interpretability features to the model by uncovering the local dominance of certain parameters on prediction outcomes. The following key conclusions can be drawn from this research work:

1. Developed five key classification intelligent models with the highest efficiency in detecting stuck pipe occurrences, which included Logistic Regression, K-Nearest Neighbors, Support Vector Machines, Decision Trees and CatBoost. These models were trained and results were validated using actual field data collected from 10 wells with 61 datasets consisted of 610 data points.
2. Two intelligent models with the highest accuracy and robust performance were selected through error evaluation analysis, which are Decision Trees and CatBoost intelligent models. DT demonstrated the highest accuracy of 99.59% at the field level, with precision 1, recall 1 and F1 score of 1. While CatBoost outperformed DT in the individual section

evaluations with accuracy of 100%, precision 1, recall 1 and F1 score of 1. LR and KNN models also demonstrated good performance with accuracy of more than 80% and precision, recall and F1 score of more than 0.8.

3. Sensitivity studies based on feature importance evaluation and SHAP analysis of CatBoost classifier identified salt content as the main influencer on stuck pipe occurrences, followed by drilling parameters such as RPM and Flow rate. SHAP evaluation indicated some local dominance of formation effect by defining the clay content and depth; while, global evaluation showed significance of mud properties of mud weight and gel strength as main contributors to stuck pipe condition.

In addition, the following recommendations are made for further studies of stuck pipe detection problem:

1. To evaluate developed intelligent models using different field data including high-frequency surface parameters to enhance the models' generalization capability.

2. To enhance the models with parameters obtained from physics-based approaches to improve prediction ability in other complex geological environments.

3. To assess potential of using unsupervised learning algorithms for enhancing performance of stuck pipe prediction.

REFERENCES

Aadnoy, B. S., & Djurhuus, J. (2008, August 25). Theory and Application of a New Generalized Model for Torque and Drag. *IADC/SPE Asia Pacific Drilling Technology Conference and Exhibition*. https://doi.org/10.2118/114684-MS

Abbas, A. K., Flori, R., Almubarak, H., Dawood, J., Abbas, H., & Alsaedi, A. (2019, September 23). Intelligent Prediction of Stuck Pipe Remediation Using Machine Learning Algorithms. *SPE Annual Technical Conference and Exhibition*. https://doi.org/10.2118/196229-MS

Ahmed, O. S., Aman, B. M., Zahrani, M. A., & Ajikobi, F. I. (2019, November 11). Stuck Pipe Early Warning System Utilizing Moving Window Machine Learning Approach. *Abu Dhabi International Petroleum Exhibition & Conference*. https://doi.org/10.2118/197674-MS

Akhmetzhanov, A., Zholtayev, G., Djeddou, A., Akhmetzhanova, G., & Oraz, B. (2020). Post-salt trapping mechanism of south-east Pre-Caspian and its application to petroleum exploration. *Journal of Petroleum Exploration and Production Technology*, *10*(7), 2645–2653. https://doi.org/10.1007/s13202-020-00971-9

Al Dushaishi, M. F., Abbas, A. K., Alsaba, M., Abbas, H., & Dawood, J. (2021). Data-driven stuck pipe prediction and remedies. *Upstream Oil and Gas Technology*, *6*. https://doi.org/10.1016/j.upstre.2020.100024

Albaiyat, I. (2012). *Implementing Artificial Neural Networks and Support Vector Machines in Stuck Pipe Prediction*. Texas Tech University.

Aljubran, M. J., Al-Yami, A. S., & Madan, M. A. (2017, March 6). Best Practices to Prevent Stuck Pipe Incidents in Offshore Shaly Formations. *SPE Middle East Oil & Gas Show and Conference*. https://doi.org/10.2118/183717-MS

Al-Malki, M., Mahmoud, A. A., & Elkatatny, S. (2023). Performance of the Hole Cleaning Factor in Predicting the Hole Cleaning Conditions in Vertical and Deviated Wells: Real Case Studies.

Arabian Journal for Science and Engineering, *48*(12), 17119–17127. https://doi.org/10.1007/s13369-023-08042-x

Alshaikh, A. A., Albassam, M. K., Al Gharbi, S. H., & Al-Yami, A. S. (2018, November 12). Detection of Stuck Pipe Early Signs and the Way Toward Automation. *Abu Dhabi International Petroleum Exhibition & Conference*. https://doi.org/10.2118/192975-MS

Alshaikh, A. A., Magana-Mora, A., Al Gharbi, S., Al-Yami, A., & Aramco, S. (2019). *IPTC-19394-MS Machine Learning for Detecting Stuck Pipe Incidents: Data Analytics and Models Evaluation*. http://onepetro.org/IPTCONF/proceedings-pdf/19IPTC/2-19IPTC/D021S021R003/1127235/iptc-19394-ms.pdf/1

Alzahrani, M., Alotaibi, B., & Aman, B. (2022). Novel Stuck Pipe Troubles Prediction Model Using Reinforcement Learning. *International Petroleum Technology Conference, IPTC 2022*. https://doi.org/10.2523/IPTC-22151-MS

Amer, A., Dearing, H., & Jones, R. (2016). *SPE-180326-MS Drilling Through Salt Formations: A Drilling Fluids Review Newpark Drilling Fluids; M. Sergiacomo, ENI.*

Antonini, A. S., Tanzola, J., Asiain, L., Ferracutti, G. R., Castro, S. M., Bjerg, E. A., & Ganuza, M. L. (2024). Machine Learning model interpretability using SHAP values: Application to Igneous Rock Classification task. *Applied Computing and Geosciences*, *23*, 100178. https://doi.org/10.1016/j.acags.2024.100178

Belaskie, J. P., Mccann, D. P., Leshikar, J. F., & Schlumberger, A. /. (1994). *IADC/SPE 27494 A Practical Method To Minimize Stuck Pipe Integrating Surface and MWD Measurements*. http://onepetro.org/SPEDC/proceedings-pdf/94DC/94DC/SPE-27494-MS/3769561/spe-27494-ms.pdf/1

Biegler, M. W. (1994). *IADC/SPE 27529 Advances in Prediction of Stuck Pipe Using Multivariate Statistical Analysis*. http://onepetro.org/SPEDC/proceedings-pdf/94DC/94DC/SPE-27529-MS/3769034/spe-27529-ms.pdf/1

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning*.

Bourel, M., & Segura, A. M. (2018). Multiclass classification methods in ecology. *Ecological Indicators*, *85*, 1012–1021. https://doi.org/10.1016/j.ecolind.2017.11.031

Bowes, C., Procter, R., Ringrose, D., Lancaster, D., Coghill, K., Roberts Earl, D., & Forex, S.-S. (1997). *997 Drillers Stuck pipe Handbook 1997 Guidelines & Drillers Handbook Credits written by Sedco Forex Project Steering & Review Team*. http://www.stuckpipe.co.uk

Bradley, W. B., Jarman, D., Plott, R. S., Wood, R. D., Schofield, T. R., Auflick, R. A., & Cocking, D. (1991, March 11). A Task Force Approach to Reducing Stuck Pipe Costs. *SPE/IADC Drilling Conference*. https://doi.org/10.2118/21999-MS

Brandon, N. P., Panesar, S. S., Bonanos, N., Fogarty, P. O., & Mahmood, M. N. (1993). The effect of cathodic currents on friction and stuck pipe release in aqueous drilling muds. *Journal of Petroleum Science and Engineering*, *10*(2), 75–82. https://doi.org/10.1016/0920-4105(93)90032-A

Brankovic, A., Matteucci, M., Restelli, M., Ferrarini, L., Piroddi, L., Spelta, A., & Zausa, F. (2021). Data-driven indicators for the detection and prediction of stuck-pipe events in oil&gas drilling operations. *Upstream Oil and Gas Technology*, *7*. https://doi.org/10.1016/j.upstre.2021.100043

Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (2017). *Classification And Regression Trees*. Routledge. https://doi.org/10.1201/9781315139470

Campbell, S., Bueno, R., Eren, S., & Abdullah, M. F. (2021, March 16). Salt Drilling: The State of the Art. *Day 1 Tue, March 23, 2021*. https://doi.org/10.2523/IPTC-21165-MS

Carcione, J. M., Helle, H. B., & Gangi, A. F. (2006). Theory of borehole stability when drilling through salt formations. *GEOPHYSICS*, *71*(3), F31–F47. https://doi.org/10.1190/1.2195447

Cayeux, E., Daireaux, B., Wolden Dvergsnes, E., & Sælevik, G. (2012). Early Symptom Detection on the Basis of Real-Time Evaluation of Downhole Conditions: Principles and Results From Several North Sea Drilling Operations. *SPE Drilling & Completion*, *27*(04), 546–558. https://doi.org/10.2118/150422-PA

Chamkalani, A., Shahri, M. P., & Poordad, S. (2013). *SPE 164003 Support Vector Machine Model: A New Methodology for Stuck Pipe Prediction*. http://onepetro.org/SPEUGM/proceedings-pdf/13UGM/All-13UGM/SPE-164003-MS/1583309/spe-164003-ms.pdf

Courteille, J. M., & Zurdo, C. (1985, September 22). A New Approach to Differential Sticking. *SPE Annual Technical Conference and Exhibition*. https://doi.org/10.2118/14244-MS

Cristianini, N., & Shawe-Taylor, J. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge University Press. https://doi.org/10.1017/CBO9780511801389

D'Amicis, S., Pagani, M., Matteucci, M., Piroddi, L., Spelta, A., & Zausa, F. (2023). Stuck pipe prediction from rare events in oil drilling operations. *Upstream Oil and Gas Technology*, *11*, 100096. https://doi.org/10.1016/j.upstre.2023.100096

D'Angelo, J., Ashok, P., van Oort, E., Shahri, M., Thetford, T., Nelson, B., Behounek, M., & White, M. (2019, March 4). Unplanned Tortuosity Index: Separating Directional Drilling Performance from Planned Well Geometry. *SPE/IADC International Drilling Conference and Exhibition*. https://doi.org/10.2118/194099-MS

Dorogush, A. V., Ershov, V., & Gulin, A. (2018). *CatBoost: gradient boosting with categorical features support*. http://arxiv.org/abs/1810.11363

Drilling Manual. (n.d.). Https://Www.Drillingmanual.Com/. Retrieved April 1, 2025, from https://www.drillingmanual.com/

Dupriest, F. E., Elks, W. C., & Ottesen, S. (2011). Design Methodology and Operational Practices Eliminate Differential Sticking. *SPE Drilling & Completion*, *26*(01), 115–123. https://doi.org/10.2118/128129-PA

Dusseault, M. B., Vincent Maury, Francesco Sanfilippo, & Frédéric J. Santarelli. (2004). *Drilling Around Salt: Risks, Stresses, And Uncertainties*.

Elahifar, B., & Hosseini, E. (2022). Machine learning algorithm for prediction of stuck pipe incidents using statistical data: case study in middle east oil fields. *Journal of Petroleum Exploration and Production Technology*, *12*(7), 2019–2045. https://doi.org/10.1007/s13202-021-01436-3

Elgibaly, A. A., Farhat, M. S., Trant, E. W., & Kelany, M. (2017). A study of friction factor model for directional wells. *Egyptian Journal of Petroleum*, *26*(2), 489–504. https://doi.org/10.1016/j.ejpe.2016.07.004

Elmousalami, H. H., & Elaskary, M. (2020). Drilling stuck pipe classification and mitigation in the Gulf of Suez oil fields using artificial intelligence. *Journal of Petroleum Exploration and Production Technology*, *10*(5), 2055–2068. https://doi.org/10.1007/s13202-020-00857-w

Falconer, I. G., Belaskie, J. P., & Variava, F. (1989, February 28). Applications of a Real Time Wellbore Friction Analysis. *SPE/IADC Drilling Conference*. https://doi.org/10.2118/18649-MS

Fawcett, T. (2006). An introduction to ROC analysis. *Pattern Recognition Letters*, *27*(8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010

Ferreira, P. A. L. A., Carvalho, D. J. L., Rodrigues, R. M., Schnell, D. M., Thomson, I. J., Baptista, R. C., & Alves, S. B. (2015). *OTC-25838-MS Automated Decision Support and Expert Collaboration Avoid Stuck Pipe and Improve Drilling Operations in Offshore Brazil Subsalt*

Well. http://onepetro.org/OTCONF/proceedings-pdf/15OTC/All-15OTC/OTC-25838-MS/1360462/otc-25838-ms.pdf/1

Gonzalez, O., Bernat, H., & Moore, P. (2007, November 11). The Extraction of Mud-Stuck Tubulars Using Vibratory Resonant Techniques. *SPE Annual Technical Conference and Exhibition*. https://doi.org/10.2118/109530-MS

Gulsrud, T. O., Nybø, R., & Bjørkevoll, K. S. (2009). *SPE 123374 Statistical Method for Detection of Poor Hole Cleaning and Stuck Pipe*. http://onepetro.org/SPEOE/proceedings-pdf/09OE/09OE/SPE-123374-MS/4178164/spe-123374-ms.pdf/1

Guzman, J. M., Khalil, M. E., Aramco, S., Orban, N., Mohiuddin, M. A., Verma, J., & Ganda, S. (2012). *the SPE Saudi Arabia Section Technical Symposium and Exhibition*. http://onepetro.org/SPESATS/proceedings-pdf/12SATS/12SATS/SPE-160875-MS/1649086/spe-160875-ms.pdf/1

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning*. Springer New York. https://doi.org/10.1007/978-0-387-84858-7

Heitmann, N., & Burgos, E. C. (2015, March 17). Freeing Differential Stuck-Pipe with Nitrogen Reduces Significantly Lost-In-Hole Drill Strings. *SPE/IADC Drilling Conference and Exhibition*. https://doi.org/10.2118/173168-MS

Helmick, W. E., & Longley A. J. (1957). *Pressure-differential Sticking of Drill Pipe and How it can be Avoided or Relieved*.

Hemmati-Sarapardeh, A., Menad Nait Amar, Aydin Larestani, & Sassan Hajirezaie. (2020). *Applications of Artificial Intelligence Techniques in the Petroleum Industry*. Elsevier. https://doi.org/10.1016/C2018-0-04421-7

Hempkins, W. B., & Lohec, W. E. (1987). *Multivariate Statistical Analysis of Stuck Drillpipe Situations*. http://onepetro.org/DC/article-pdf/2/03/237/2640848/spe-14181-pa.pdf/1

Hilfiger, M. G., Thaemlitz, C. J., & Moellendick, E. (2017, March 6). Advanced Spotting Fluids for Remediation of Differentially Stuck Pipe. *SPE Middle East Oil & Gas Show and Conference*. https://doi.org/10.2118/183959-MS

Hosmer, D. W., & Lemeshow, S. (2000). *Applied Logistic Regression*. Wiley. https://doi.org/10.1002/0471722146

Howard, J. A., & Glover, S. B. (1994, February 15). Tracking Stuck Pipe Probability While Drilling. *IADC/SPE Drilling Conference*. https://doi.org/10.2118/27528-MS

Hunter, D., Baroid, NL., & Adams, N. (1978, May 7). Laboratory And Field Data Indicate Water Base Drilling Fluids That Resist Differential-Pressure Pipe Sticking. *All Days*. https://doi.org/10.4043/3239-MS

Inoue, T., Nakagawa, Y., Yokosuka, J., Wada, J. R., Miyoshi, K., Abe, S., Tokyo, J., Kouhei Kuroda, J., Tokyo, J., Tamamura, J. H., Tokyo, I., & Hakan Bilen, J. (2021). *ATTEMPT OF EARLY STUCK DETECTION USING UNSUPERVISED DEEP LEARNING WITH PROBABILITY MIXTURE MODEL*. http://asmedigitalcollection.asme.org/OMAE/proceedings-pdf/OMAE2021/85208/V010T11A004/6781078/v010t11a004-omae2021-62739.pdf

Isambourg, P., Ottesen, S., Benaissa, S., & Marti, J. (1999, March 9). Down-Hole Simulation Cell for Measurement of Lubricity and Differential Pressure Sticking. *SPE/IADC Drilling Conference*. https://doi.org/10.2118/52816-MS

Issa, M. A., A. Alrazzaq, A. A. A., & Mukhtar, Y. (2023). Review of the Mechanisms for Preventing, Diagnosing, and Treatment of Pipe Sticking in Drilling Operations. *Iraqi Journal of Chemical and Petroleum Engineering*, *24*(3), 133–140. https://doi.org/10.31699/ijcpe.2023.3.13

Jahanbakhshi, R., & Keshavarzi, R. (2016). Intelligent Classifier Approach for Prediction and Sensitivity Analysis of Differential Pipe Sticking: A Comparative Study. *Journal of Energy Resources Technology, Transactions of the ASME*, *138*(5). https://doi.org/10.1115/1.4032831

Jahanbakhshi, R., Keshavarzi, R., Researchers, Y., Club, E., Branch, R., & Shoorehdeli, A. (2012). *Intelligent Prediction of Differential Pipe Sticking by Support Vector Machine Compared With Conventional Artificial Neural Networks: An Example of Iranian Offshore Oil Fields*. http://onepetro.org/DC/article-pdf/27/04/586/2092829/spe-163062-pa.pdf/1

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An Introduction to Statistical Learning*. Springer US. https://doi.org/10.1007/978-1-0716-1418-1

Jardine, S. I., Mccann, D. P., Barber, S. S., & Schlumberger, S. F. (1992). *This paper was prepared for presentation at the 1992 IADCISPE Drilling Conference*. http://onepetro.org/SPEDC/proceedings-pdf/92DC/92DC/SPE-23915-MS/3731245/spe-23915-ms.pdf/1

Johancsik, C. A., Friesen, D. B., & Dawson, R. (1984). Torque and Drag in Directional Wells-Prediction and Measurement. *Journal of Petroleum Technology*, *36*(06), 987–992. https://doi.org/10.2118/11380-PA

Kaneko, T., Inoue, T., Nakagawa, Y., Wada, R., Abe, S., Yasutake, G., & Fujita, K. (2024). Hybrid Approach Using Physical Insights and Data Science for Stuck-Pipe Prediction. In *SPE Journal* (Vol. 641). http://onepetro.org/SJ/article-pdf/29/02/641/3369014/spe-218013-pa.pdf

Kaneko, T., Inoue, T., Nakagawa, Y., Wada, R., Miyoshi, K., Abe, S., Kuroda, K., & Fujita, K. (2023). Hybrid Approach Using Physical Insights and Data Science for Early Stuck Detection. *Proceedings of the Annual Offshore Technology Conference*, *2023-May*. https://doi.org/10.4043/32532-MS

Katzmann, L., Mayer, H., Arévalo, P., & Forshaw, M. (2024). Machine Learning Accelerates and Facilitates the Prediction of Stuck Pipe Events. *Society of Petroleum Engineers - ADIPEC 2024*. https://doi.org/10.2118/222438-MS

Lesage, M., Falconer, I. G., & Wick, C. J. (1988). Evaluating Drilling Practice in Deviated Wells With Torque and Weight Data. *SPE Drilling Engineering*, *3*(03), 248–252. https://doi.org/10.2118/16114-PA

Li, Z. (2022). Extracting spatial effects from machine learning model using local interpretation method: An example of SHAP and XGBoost. *Computers, Environment and Urban Systems*, *96*, 101845. https://doi.org/10.1016/j.compenvurbsys.2022.101845

Liu, M., Song, X., Zhu, Z., Li, G., Xiao, H., Fu, L., Pan, T., Li, X., & Yang, Y. (2024, September 20). Inertial Effect of Drill Pipe Friction Trend and Early Warning of Stuck Pipe. *SPE Annual Technical Conference and Exhibition*. https://doi.org/10.2118/220991-MS

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S.-I. (2019). *Explainable AI for Trees: From Local Explanations to Global Understanding*. http://arxiv.org/abs/1905.04610

Magana-Mora, A., Gharbi, S., Alshaikh, A., & Al-Yami, A. (2019). *SPE-194980-MS AccuPipePred: A Framework for the Accurate and Early Detection of Stuck Pipe for Real-Time Drilling Operations*. http://onepetro.org/SPEMEOS/proceedings-pdf/19MEOS/19MEOS/D032S077R003/4030084/spe-194980-ms.pdf/1

Mahmood, A. T. K., & Assi, A. H. (2024). A Comprehensive Review of the Pipe Sticking Mechanism in Oil Well Drilling Operations. *Journal of Engineering*, *30*(11), 50–70. https://doi.org/10.31026/j.eng.2024.11.04

Mal, A., Ødegård, S. I., Helgeland, S., Sinaga, S. Z., & Svendsen, M. (2022). Prediction of Stuck Pipe Incidents Using Models Powered by Deep Learning and Machine Learning. *SPE - International Association of Drilling Contractors Drilling Conference Proceedings*, *2022-March*. https://doi.org/10.2118/208778-MS

Malki, M. A., Abughaban, M. F., Alshawabkeh, A. A., & Guimaraes, T. T. (2024). Real-Time Detection of Stuck Pipe Utilizing Hybrid AI-Physical Prediction Models. *International Petroleum Technology Conference, IPTC 2024*. https://doi.org/10.2523/IPTC-24078-MS

Mats Håpnes. (2014). *Drilling in Salt Formations and Rate of Penetration Modelling*. Norwegian University of Science and Technology.

Meor Hashim, M. M., Yusoff, M. H., Arriffin, M. F., Mohamad, A., Gomes, D., Jose, M., & Tengku Bidin, T. E. (2021, March 16). Utilizing Artificial Neural Network for Real-Time Prediction of Differential Sticking Symptoms. *Day 7 Mon, March 29, 2021*. https://doi.org/10.2523/IPTC-21221-MS

Miri, R., Sampaio, J., Afshar, M., & Lourenco, A. (2007, June 27). Development of Artificial Neural Networks To Predict Differential Pipe Sticking in Iranian Offshore Oil Fields. *International Oil Conference and Exhibition in Mexico*. https://doi.org/10.2118/108500-MS

Mitchell. (2001). *Trouble-Free Drilling: Stuck Pipe Prevention* (3rd ed.).

Mitchell, R. F., & Samuel, R. (2009). How Good Is the Torque/Drag Model? *SPE Drilling & Completion*, *24*(01), 62–71. https://doi.org/10.2118/105068-PA

Montes, A. C., Ashok, P., & van Oort, E. (2023). Stuck Pipe Prediction in Utah FORGE Geothermal Wells. *Proceedings - SPE Annual Technical Conference and Exhibition*, *2023-October*. https://doi.org/10.2118/214783-MS

Mopuri, K. R., Bilen, H., Tsuchihashi, N., Wada, R., Inoue, T., Kusanagi, K., Nishiyama, T., & Tamamura, H. (2022). Early sign detection for the stuck pipe scenarios using unsupervised deep learning. *Journal of Petroleum Science and Engineering*, *208*. https://doi.org/10.1016/j.petrol.2021.109489

Muqeem, M. A., Weekse, A. E., & Al-Hajji, A. A. (2012, April 8). Stuck Pipe Best Practices – A Challenging Approach to Reducing Stuck Pipe Costs. *SPE Saudi Arabia Section Technical Symposium and Exhibition*. https://doi.org/10.2118/160845-MS

Murillo, A., Neuman, J., & Samuel, R. (2009). *SPE 120128 Pipe Sticking Prediction and Avoidance Using Adaptive Fuzzy Logic and Neural Network Modeling*. http://onepetro.org/SPEOKOG/proceedings-pdf/09POS/All-09POS/SPE-120128-MS/2719143/spe-120128-ms.pdf/1

Murtaza, M., Kamal, M. S., Hussain, S. M. S., Mahmoud, M., & Syed, N. A. (2020). Quaternary ammonium gemini surfactants having different spacer length as clay swelling inhibitors: Mechanism and performance evaluation. *Journal of Molecular Liquids*, *308*, 113054. https://doi.org/10.1016/j.molliq.2020.113054

Naraghi, M. E., & Jamshidi, S. (2013). *Adaptive Neuro Fuzzy Inference System and Artificial Neural Networks: reliable approaches for pipe stuck prediction Article in AUSTRALIAN JOURNAL OF BASIC AND APPLIED SCIENCES · January 2013 CITATIONS 6 READS 200*. https://www.researchgate.net/publication/277331902

Nezhad, M. M., Ashoori, S., Hooshmand, P., & Mirzaee, M. (2012). Stuck Drill Pipe Prediction with Networks Neural in Maroon Field. *J. Basic. Appl. Sci. Res*, *2*(6), 5570–5575. www.textroad.com

Omojuwa, E. O., Osisanya, S. O., & Ahmed, R. (2011, July 30). Properties of Salt Formations Essential for Modeling Instabilities While Drilling. *Nigeria Annual International Conference and Exhibition*. https://doi.org/10.2118/150801-MS

Othman, E. Bin, Gomes, D., Bidin, T. E. B. T., Hashim, M. M. H. M., Yusoff, M. H., Arriffin, M. F., & Ghazali, R. (2022). Application of Machine Learning to Augment Wellbore Geometry-Related Stuck Pipe Risk Identification in Real Time. *Offshore Technology Conference Asia, OTCA 2022*. https://doi.org/10.4043/31695-MS

Outmans, H. D. (1958). Mechanics of Differential Pressure Sticking of Drill Collars. *Transactions of the AIME*, *213*(01), 265–274. https://doi.org/10.2118/963-G

Payrazyan, V. K., & Robinson, T. S. (2023). Leveraging Targeted Machine Learning for Early Warning and Prevention of Stuck Pipe, Tight Holes, Pack Offs, Hole Cleaning Issues and Other Potential Drilling Hazards. *Proceedings of the Annual Offshore Technology Conference*, *2023-May*. https://doi.org/10.4043/32169-MS

Reid, P. I., Meeten, G. H., Way, P. W., Clark, P., Chambers, B. D., Gilmour, A., & Sanders, M. W. (2000). Differential-Sticking Mechanisms and a Simple Wellsite Test for Monitoring and Optimizing Drilling Mud Properties. *SPE Drilling & Completion*, *15*(02), 97–104. https://doi.org/10.2118/64114-PA

Rostami, H., & Manshad, A. K. (2014). A new support vector machine and artificial neural networks for prediction of stuck pipe in drilling of oil fields. *Journal of Energy Resources Technology, Transactions of the ASME*, *136*(2). https://doi.org/10.1115/1.4026917

Salminen, K., Cheatham, C., Smith, M., & Valiulin, K. (2016). *IADC/SPE-178888-MS Stuck Pipe Prediction Using Automated Real-Time Modeling and Data Analysis*. http://onepetro.org/SPEDC/proceedings-pdf/16DC/16DC/D031S026R002/3791844/spe-178888-ms.pdf/1

Samuel, R. (2010). Friction factors: What are they for torque, drag, vibration, bottom hole assembly and transient surge/swab analyses? *Journal of Petroleum Science and Engineering*, *73*(3–4), 258–266. https://doi.org/10.1016/j.petrol.2010.07.007

Saporetti, C. M., da Fonseca, L. G., Pereira, E., & de Oliveira, L. C. (2018). Machine learning approaches for petrographic classification of carbonate-siliciclastic rocks using well logs and textural information. *Journal of Applied Geophysics*, *155*, 217–225. https://doi.org/10.1016/j.jappgeo.2018.06.012

Sewell, M. (2011). *Ensemble Learning.*

Shadizadeh, S. R., Karimi, F., & Zoveidavianpoor, M. (2010). Drilling Stuck Pipe Prediction in Iranian Oil Fields: An Artificial Neural Network Approach. In *Iranian Journal of Chemical Engineering* (Vol. 7, Issue 4).

Shahbazi, K., & Shahri, M. P. (2011). The development of a window for stuck pipe prediction. *Petroleum Science and Technology*, *30*(2), 176–192. https://doi.org/10.1080/10916461003792302

Shapley, L. S. (1953). 17. A Value for n-Person Games. In *Contributions to the Theory of Games (AM-28), Volume II* (pp. 307–318). Princeton University Press. https://doi.org/10.1515/9781400881970-018

Sheppard, M. C., Wick, C., & Burgess, T. (1987). Designing Well Paths to Reduce Drag and Torque. *SPE Drilling Engineering*, *2*(04), 344–350. https://doi.org/10.2118/15463-PA

Shoraka, S. A., Shadizadeh, S. R., & Pordel Shahri, M. (2011, July 30). Prediction of Stuck Pipe in Iranian South Oil Fields Using Multivariate Statistical Analysis. *Nigeria Annual International Conference and Exhibition*. https://doi.org/10.2118/151076-MS

Singh, A., Yadav, A., & Rana, A. (2013). K-means with Three different Distance Metrics. *International Journal of Computer Applications*, *67*(10), 13–17. https://doi.org/10.5120/11430-6785

Siruvuri, C., Digital, H., Solutions, C., Nagarakanti, S., Industries, N., & Samuel, R. (2006). *Stuck Pipe Prediction and Avoidance: A Convolutional Neural Network Approach*. http://onepetro.org/SPEDC/proceedings-pdf/06DC/06DC/SPE-98378-MS/3775333/spe-98378-ms.pdf/1

Tinseau, E., Bartier, D., Hassouta, L., Devol-Brown, I., & Stammose, D. (2006). Mineralogical characterization of the Tournemire argillite after in situ interaction with concretes. *Waste Management*, *26*(7), 789–800. https://doi.org/10.1016/j.wasman.2006.01.024

Tsuchihashi, N., Wada, R., Ozaki, M., Inoue, T., Mopuri, K. R., Bilen, H., Nishiyama, T., Fujita, K., & Kusanagi, K. (2021). Early Stuck Pipe Sign Detection with Depth-Domain 3D Convolutional Neural Network Using Actual Drilling Data. *SPE Journal*, *26*(02), 551–562. https://doi.org/10.2118/204462-pa

Vaferi, B., Dehbashi, M., Alibak, A. H., & Yousefzadeh, R. (2024). Exploring the performance of machine learning models to predict carbon monoxide solubility in underground pure/saline water. *Marine and Petroleum Geology*, *162*, 106742. https://doi.org/10.1016/j.marpetgeo.2024.106742

Vapnik, V., & Cortes, C. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. https://doi.org/10.1007/BF00994018

Wang, H., & Samuel, R. (2016). 3D Geomechanical Modeling of Salt-Creep Behavior on Wellbore Casing for Presalt Reservoirs. *SPE Drilling & Completion*, *31*(04), 261–272. https://doi.org/10.2118/166144-PA

Weakley, R. R. (1990). *SPE SPE 20410 Use of Stuck Pipe Statistics To Reduce the Occurrence of Stuck Pipe*. http://onepetro.org/SPEATCE/proceedings-pdf/90SPE/90SPE/SPE-20410-MS/3472181/spe-20410-ms.pdf/1

Willersrud, A. (2015). *Model-Based Diagnosis of Drilling Incidents*.

Wu, X., Kumar, V., Ross Quinlan, J., Ghosh, J., Yang, Q., Motoda, H., McLachlan, G. J., Ng, A., Liu, B., Yu, P. S., Zhou, Z.-H., Steinbach, M., Hand, D. J., & Steinberg, D. (2008). Top 10 algorithms in data mining. *Knowledge and Information Systems*, *14*(1), 1–37. https://doi.org/10.1007/s10115-007-0114-2

Yousefzadeh, R., Kazemi, A., & Al-Maamari, R. S. (2024). Application of power-law committee machine to combine five machine learning algorithms for enhanced oil recovery screening. *Scientific Reports*, *14*(1), 9200. https://doi.org/10.1038/s41598-024-59387-8

Zhang, F., University, Y., Islam, A., Asa, E., Zeng, H., Chen, Z., Zeng, Y., Wang, X., & Li, S. (2019). *SPE-197167-MS Real Time Stuck Pipe Prediction by Using a Combination of Physics-Based Model and Data Analytics Approach*. http://onepetro.org/SPEADIP/proceedings-pdf/19ADIP/2-19ADIP/D022S173R001/1119937/spe-197167-ms.pdf

Zhao, J., Shen, Y., Chen, W., Zhang, Z., & Johnston, S. (2017). *SPE-187512-MS Machine Learning-Based Trigger Detection of Drilling Events Based on Drilling Data* (Vol. 2). http://onepetro.org/SPEERM/proceedings-pdf/17ERM/17ERM/D023S001R005/1292541/spe-187512-ms.pdf/1

Zhu, Q., Wang, Z., Huang, J., & Bohai, C. (2019). *SPE-198672-MS Stuck Pipe Incidents Prediction Based On Data Analysis*. http://onepetro.org/SPEGOTS/proceedings-pdf/19GOTS/2-19GOTS/D022S027R001/1181930/spe-198672-ms.pdf/1

Zhu, S., Song, X., Zhu, Z., Yao, X., & Liu, M. (2022). *Intelligent Prediction of Stuck Pipe Using Combined Data-Driven and Knowledge-Driven Model*. https://doi.org/10.3390/app