RESEARCH ARTICLE

A Machine Learning Framework for Artificial Lift Method Selection with Physics-Informed Data Balancing

Sohail Nawab[1,3,*], Muhammad Ali[2], Imran Ahmed Hullio[3], Noshad[4], Ning Liu[1] and Sarfraz A. Jokhio[3]

[1] Department of Petroleum Engineering, University of Louisiana at Lafayette, Lafayette, LA 70504, United States
[2] Department of Engineering Data Science, University of Houston, Houston, TX 77204, United States
[3] Institute of Petroleum and Natural Gas Engineering, Mehran University of Engineering and Technology, Jamshoro, Sindh 76062, Pakistan
[4] College of Petroleum Engineering, China University of Petroleum, Beijing, Beijing 102249, China

Abstract

The selection of optimal artificial lift methods using machine learning remains challenging due to complex interactions among reservoir characteristics, fluid properties, and operational constraints. Conventional approaches rely on engineering expertise and static screening criteria, often insufficient to capture multifactorial dependencies. This study presents a framework for classifying the most suitable lift method from four common techniques: ESP, Gas Lift, Rod Pumps, and PCP. A dataset of 990 wells with twelve physically meaningful parameters was compiled, including depth, temperature, GOR, API gravity, reservoir pressure, water cut, production rate, viscosity, sand production, deviation, H_2S presence, and formation type. To address class imbalance, SMOTE and the proposed DI-SDG method—which derives feature-specific perturbation limits from published intra-class variability data—were evaluated. Six ML models were trained using five-fold stratified cross-validation. Random Forest achieved the best test performance (accuracy: 91.41%, precision: 92.47%, recall: 92.67%, macro-F_1: 0.9255), with XGBoost (F_1: 0.9125) and Gradient Boosting (F_1: 0.9006) also performing well. Generalization was validated via blind well testing using independent field cases. Analysis of 18 misclassified samples showed most errors occurred in overlapping operating envelopes, particularly between ESP and Gas Lift in intermediate GOR ranges. Misclassified samples averaged 64% confidence versus 82% for correct classifications, suggesting a 75% threshold for cases requiring additional evaluation. Overall, the physics-informed balancing approach provides an accurate, interpretable framework for artificial lift selection with reliable field-data performance.

Keywords: artificial lift selection, machine learning, screening criteria, class imbalance, multi-class classification, ESP, gas lift, rod pump, PCP, domain-informed data augmentation.

Submitted: 27 April 2026
Accepted: 10 June 2026
Published: 19 June 2026

Vol. 2, **No.** 3, 2026.
10.62762/RS.2026.704585

***Corresponding author:**
✉ Sohail Nawab
sohailnawabkh@hotmail.com

Citation
Nawab, S., Ali, M., Hulio, I. A., Noshad, Liu, N., & Jokhio, S. A. (2026). A Machine Learning Framework for Artificial Lift Method Selection with Physics-Informed Data Balancing. *Reservoir Science*, 2(3), 228–260.

Abbreviations and Acronyms

Symbol	Definition
AI	Artificial Intelligence
AL / ALM	Artificial Lift / Artificial Lift Method
API	American Petroleum Institute
BHT	Bottomhole Temperature
BPD	Barrels Per Day
cP	Centipoise
CV	Coefficient of Variation
DI-SDG	Domain-Informed Synthetic Data Generation
EOR / IOR	Enhanced Oil Recovery / Improved Oil Recovery
ESP	Electric Submersible Pump
GOR	Gas-Oil Ratio
HNBR	Hydrogenated Nitrile Butadiene Rubber
IPR	Inflow Performance Relationship
KNN	K-Nearest Neighbors
ML	Machine Learning
NACE	National Association of Corrosion Engineers
PCA	Principal Component Analysis
PCP	Progressive Cavity Pump
PI	Productivity Index
PVT	Pressure-Volume-Temperature
RF	Random Forest
scf/STB	Standard Cubic Feet per Stock Tank Barrel
SHAP	SHapley Additive exPlanations
SMOTE	Synthetic Minority Over-sampling Technique
SSCC	Sulfide Stress Corrosion Cracking
STB	Stock Tank Barrel
STB/day	Stock Tank Barrels per Day
SVM	Support Vector Machine
XAI	Explainable Artificial Intelligence
XGBoost	Extreme Gradient Boosting

Mathematical Symbols

Symbol	Definition
$\alpha, \alpha_0, \alpha_j$	Relative perturbation scale parameter (Mode A)
β, β_j	Absolute perturbation scale parameter (Mode B)
χ^2	Chi-square test statistic
ε_j	Zero-mean Gaussian perturbation for feature j
σ_j	Standard deviation of perturbation for feature j
L_j^c	Lower class-specific physical constraint bound for feature j, class c
U_j^c	Upper class-specific physical constraint bound for feature j, class c
x_j^{seed}	j-th feature value of the seed sample drawn from class c
$\tilde{x}_j$	Perturbed (synthetic) value of feature j
X_c	Set of all training samples belonging to class c
X_c'	Augmented training set for class c
N_t	Target sample count per class after augmentation
Ω	Set of all artificial lift classes
d	Number of input features ($d = 12$)
c	Artificial lift class index
k	Number of nearest neighbours (SMOTE parameter)
p_{flip}	Bernoulli flip probability for categorical features ($= 0.20$)
P_r	Reservoir pressure (psia)
P_{wf}	Flowing bottomhole pressure (psia)
Q	Production rate (STB/day)
r	Pearson correlation coefficient
ΔF_1	Difference in macro-averaged F1-score between two configurations
μ	Mean prediction confidence score

Units

Unit	Definition
API	API gravity (dimensionless index of crude oil density)
F	Degrees Fahrenheit
cP	Centipoise (dynamic viscosity)
ft	Feet
psi / psia	Pounds per square inch / absolute
scf/STB	Standard cubic feet per stock tank barrel
STB/day	Stock tank barrels per day

1 Introduction

The selection of an artificial lift method (ALM) is an important engineering decisions that affects both the short-term and the long-term operating conditions and economics of an oil well [1]. Artificial lift becomes crucial when reservoir pressure falls below the flowing bottomhole pressure, making natural flow of hydrocarbons impossible. A significant difference between selecting an artificial lift method for an existing well and a new well is in historically available production data. In existing wells, full production history over the well's life cycle can be used to support this selection process. A number of operational and reservoir-related factors must be taken into account when selecting an artificial lift method [2]. These factors include, but are not limited to, fluid properties and reservoir characteristics. For example, if a well has a high gas–oil ratio (GOR), the lift system must be capable of handling free gas effectively. In addition, when dealing with higher-viscosity fluids, there is an increased need for larger rod strings and/or larger pumps in order to maintain an adequate production rate. Furthermore, if the bottomhole temperature is 300 °F (149 °C) or above, specialized materials will likely be required in order to withstand corrosive conditions at elevated temperatures [2].

Historically, the selection of an artificial lift method has followed one of two approaches: reliance on field expertise or the application of heuristic screening guidelines [3]. Engineers would select the artificial lift system based on prior experience and apply analogous reasoning derived from comparable well conditions. In both cases, however, limitations exist; neither approach provides a systematic basis for artificial lift selection in novel or atypical well environments, nor do they account for the full range of parameters that influence lift method selection [4].

More recently, machine learning has been proposed as a decision-support tool for artificial lift system selection [5]. ML algorithms can process large datasets

and capture complex, non-linearities among various input variables [6]. This makes it possible to identify multivariate combinations associated with optimal lift system performance. In addition, these models can be retrained as the new field data becomes available, making them practical by changing reservoir and operational conditions.

Recent literature confirms that adoption of machine learning within energy industry, specifically petroleum, is accelerating. The relevant applications in the artificial lift domain include optimization of equipment placement and sizing, prediction of gas lift production rates, ESP predictive maintenance, and well integrity diagnostics [4, 7–10]. Former studies have mostly addressed different and isolated aspects of the artificial lift problem. For instance, Mahdi et al. [7] developed an artificial lift selection model using field production data from Sudan, identifying cumulative produced fluid, wellhead pressure, gas–oil ratio, produced water, gas production rate, and the implementation of Enhanced Oil Recovery/Improved Oil Recovery (EOR/IOR) operations as the most influential selection factors. More recent contributions have incorporated model interpretability techniques such as SHAP (SHapley Additive exPlanations) values; however, most of this body of work revolves on operational optimization rather than the systematic selection of lift method type.

For instance, a study published in Scientific Reports [10] applied SHAP analysis solely to the prediction of oil production rates in gas lift wells. Similarly, a study published in the Journal of Petroleum Exploration and Production Technology (JPET) [9] used interpretable ensemble learning models for predictive maintenance and failure classification in ESP systems.

One of the most exciting areas for machine learning advancement is related to optimization of artificial lift technologies. Unfortunately, one of the largest challenges facing machine learning in the artificial lift space is that there is a significant class imbalance between different lift types represented in production datasets. For example, ESPs and Gas Lifts are represented many times more often than Rod Pumps or PCPs, which can impact the way models are trained, leading to lower predictive performance specifically for underrepresented lift technologies. Moreover, most existing studies have focused on improving predictive accuracy using field production data, with less attention given to physics-constrained data balancing approaches or interpretable explanations of model decisions. As a result, there remains a gap for an explainable artificial lift selection framework that maintains its standard under imbalanced conditions, preserves engineering practicality during data augmentation, and provides transparent justification for recommended lift selections.

This paper addresses the identified research gap by developing a domain-informed ML framework for multi-class artificial lift method selection. A proposed novel Domain-Informed Synthetic Data Generation (DI-SDG) approach is used to generate synthetic training samples that reflect field actual production characteristics while remaining within physically and operationally engineering limits. Unlike conventional imbalance handling techniques such as SMOTE, DI-SDG integrates established engineering screening criteria directly into the data generation process [11, 12]. SHAP values are then used to assess and interpret model decisions in relation to established petroleum engineering principles. The principal contributions of this study are summarized as follows:

1. A dataset curated of 990 real wells representing four major artificial lift methods was made using twelve domain-informed input variables.

2. A novel synthetic data generation method, Domain-Informed Synthetic Data Generation (DI-SDG), is proposed to address class imbalance by constraining all synthetic samples within established field screening limits.

3. The performance of DI-SDG was benchmarked against the Synthetic Minority Over-sampling Technique (SMOTE) across six machine learning algorithms under identical experimental conditions.

4. Model performance was assessed using strict evaluation structure comprising hyperparameter optimization, stratified k-fold cross-validation, and a fully withheld real-well test set.

5. A transparent basis for lift method selection is established through SHAP-based feature importance and feature interaction analysis.

6. The best performing model was used and deployed as a prototype decision-support tool to facilitate field implementation.

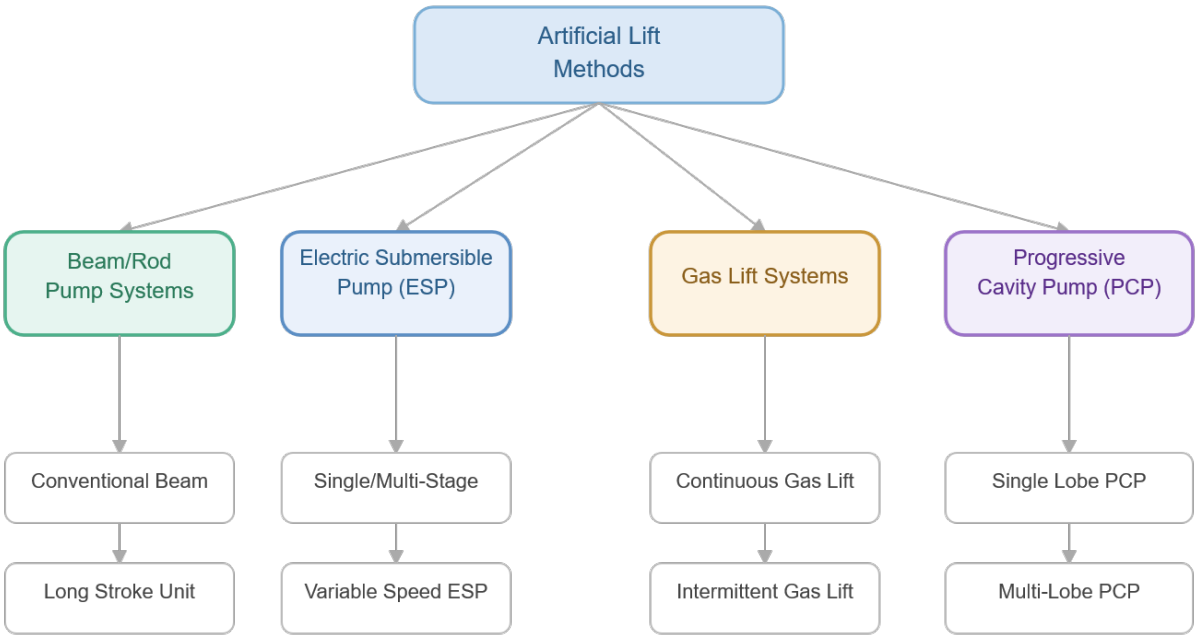

Figure 1. Classification of artificial lift methods.

2 Classification of Artificial Lift Methods

Artificial lift methods can be broadly classified into four major categories based on their operating principles and applications, as illustrated in Figure 1 [1, 2].

2.1 Electric Submersible Pump (ESP)

The downhole centrifugal pumps driven by an electric motor are called Electric Submersible Pumps (ESPs). ESPs can deliver high production rates of up to 30,000 STB/day and are widely deployed in both onshore and offshore fields. These systems are highly capable and perform well in producing fluids of moderate viscosity. Standard ESP configurations can handle moderate gas and sand levels, while pump and completion designs can be modified for higher gas and sand conditions.

2.2 Gas Lift

Gas lift is also an artificial lift method in which compressed gas is injected into the production tubing to reduce the hydrostatic pressure of the fluid column, thus enabling reservoir pressure to drive fluids to the surface [13]. This method is highly suited for wells exhibiting high gas–oil ratios (GOR), deviated wellbores, and offshore platform installations where gas compression infrastructure is available. This method also offers significant operational flexibility and can accommodate a broad range of production rates.

2.3 Rod Pump (Sucker Rod Pump)

The rod pump represents the most widely deployed artificial lift system worldwide, and is especially widespread in low-rate ("stripper") wells and shallow-to-moderate-depth reservoirs [14]. A rod pump comprises a surface-driven reciprocating prime mover connected to a downhole plunger pump via a sucker rod string. Rod pumps are generally economical for low-to-moderate production rates and can accommodate a range of fluid viscosities; however, their applicability is constrained by well depth and wellbore deviation angle.

2.4 Progressive Cavity Pumps (PCPs)

A Progressive Cavity Pump (PCP) is a positive-displacement pump that operates through the interaction of a helical rotor rotating within a double-helix elastomeric stator, that generates a continuous, non-pulsating flow. PCPs are well-suited for a wide range of demanding production environments, including heavy oil production involving high-viscosity fluids and wells producing substantial sand volumes. Their gentle fluid-handling characteristics also make them suitable

for shear-sensitive emulsions. Most PCP installations worldwide are in shallow-to-moderate-depth wells, and their adoption in heavy oil field development has increased gradually in recent years.

3 Methodology

Figure 2 presents an overview of the workflow used in this study. The methodology includes data acquisition, data cleaning and preprocessing, exploratory data analysis, feature engineering, synthetic data augmentation, model development and training, hyperparameter tuning, and model evaluation.

3.1 Artificial Lift Screening Criteria

To ensure study application to the real-world engineering practice, the dataset was anchored on 990 actual well records obtained from a field production database. Established screening criteria, referred to here as operational guardrails, were adopted as defined in Table 1. These criteria help distinguish technically feasible synthetic samples from those that are only mathematically valid but not physically realistic. The operational envelope boundaries in Table 1 were derived from decades of established industry practice and published standards [1, 2, 13–17]. Incorporating these screening criteria during synthetic data generation ensured that all augmented samples remained within physically realistic and operationally valid parameter ranges [1, 2, 15].

3.2 Data Source and Compilation

The dataset used in this study consists of 990 well records representing both current and historical operational conditions across various geological formations and production environments. The following represent the class distribution of artificial lift methods in the compiled dataset.

There is a total record of 400 Electric Submersible Pump (ESP) wells; 300 wells are operating using a Gas Lift method; 200 wells are using a Rod Pump method; and 90 wells utilise Progressive Cavity Pumps (PCP). The datasets of these different installations follow typical field deployment patterns, with ESP and Gas Lift systems being more widely represented, while substantially fewer Rod Pump and PCP installations are present in the dataset.

The information recorded for each well included twelve input features. These included well depth (ft), bottomhole temperature (°F), gas–oil ratio (GOR, scf/STB), API gravity (°API), reservoir pressure (psi), water cut (%), production rate (STB/day), fluid viscosity (cP), sand production (categorical: yes/no), wellbore deviation angle (°), H_2S presence (categorical: yes/no), and formation type (categorical: sandstone, carbonate, or shale).

The selection of these input features was based on a review of the published literature supporting their association with artificial lift method suitability [1, 2, 15, 16].

3.3 Data Cleaning and Preprocessing

Several data quality checks were applied prior to model development, as follows:

1. Identification and removal of physically inadmissible values, including negative fluid viscosity and temperatures below ambient surface conditions.
2. Validation of all feature values against the established screening criteria defined for each artificial lift method (Table 1).
3. Cross-validation of inter-feature relationships, including verification that depth–temperature pairs are consistent with expected geothermal gradient trends.

The categorical variables (formation type, sand production, and H_2S presence) were encoded using one-hot encoding. All continuous numerical features were standardized using a *StandardScaler*, yielding a mean of 0 and a standard deviation of 1, thereby ensuring consistent feature scaling across all models [18].

3.4 Physical Significance of Selected Input Features

The study included twelve input features that were selected on the basis of established petroleum engineering research and well-documented artificial lift screening practice [1, 2, 13, 14]. Each of these features represents a distinct physical or operational property that influences the feasibility, efficiency, and operational longevity of a given lift method. These features encompass five engineering dimensions relevant to lift selection: (i) well geometry (deviation, depth); (ii) fluid effects on pump performance (API gravity, viscosity, GOR); (iii) reservoir energy and deliverability (water cut, production rate, temperature, reservoir pressure); (iv) environmental and material constraints (sand production, H_2S presence); and (v) formation-specific production behavior (formation

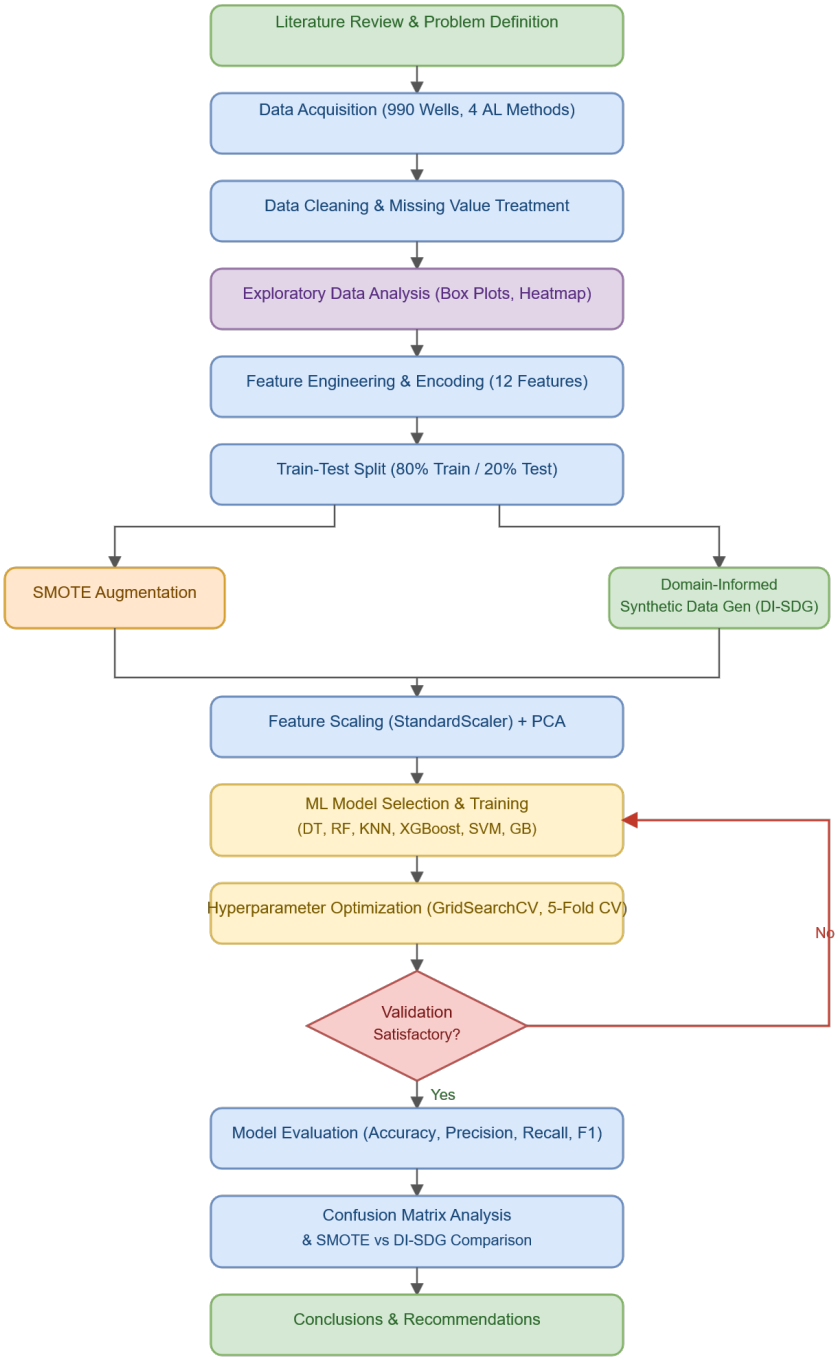

Figure 2. Research methodology flowchart.

Table 1. Screening criteria for ALM selection based on reservoir and fluid properties [1, 2, 13–17].

Parameter	ESP	Gas Lift	Rod Pump	PCP
Operating Depth (ft)	1,000–15,000(Avg. 7,500)	1,000–16,000(Avg. 8,500)	100–14,000(Avg. 4,500)	500–6,000(Avg. 3,500)
Temperature (°F)	100–300(Avg. 180)	100–350(Avg. 200)	75–250(Avg. 140)	75–180(Avg. 125)
GOR (scf/stb)	<4,000(Avg. 400)	200–Unlimited(Avg. 2,000)	2,000(Avg. 200)	500(Avg. 100)
API Gravity (°)	10–45(Avg. 30)	15–50(Avg. 32)	8–45(Avg. 28)	6–30(Avg. 18)
Reservoir Pressure (psi)	500–7,000	800–8,000	300–4,500	200–3,500
Water Cut (%)	0–95	0–90	0–98	0–98
Production Rate (bbl/day)	200–30,000(Avg. 1,500)	50–50,000(Avg. 800)	5–5,000(Avg. 100)	5–4,500(Avg. 75)
Viscosity (cp)	200(Avg. 10)	100(Avg. 5)	500(Avg. 15)	1–10,000+(Avg. 500)
Sand Production	Moderate withspecial design	Goodtolerance	Poor tomoderate	Excellenttolerance
Well Deviation (°)	0–90(with flex shaft)	0–90(Ideal for deviated)	0–30(Limited)	0–60(Moderate)

type). The underlying physical mechanism, critical applicability thresholds, associated failure modes, and discriminating role within the multi-class classifier are discussed for each of these features in the subsections below.

3.4.1 Well Depth

Well depth is one of the main screening parameters because it determines both the mechanical loads imposed on downhole equipment and the energy required by the lift system to overcome the hydrostatic head. It also governs the ambient temperature–pressure regime of the completion. For rod pump systems, the tensile and fatigue strength of the sucker rod string represents the primary limiting factor. At greater well depths, the stress imposed on the rods approaches the maximum allowable limits for API Grade D rods, and stroke efficiency decreases due to rod stretch and elastic deformation [2, 13]. Although ESPs can be deployed at depths exceeding 12,000–15,000 ft, thrust bearing load capacity, required shaft power, and cable insulation thermal demands all increase with depth [14, 19, 20]. Although Gas Lift is comparatively less depth-limited in principle, as it relies on gas injection, higher surface operating pressures are required to achieve sufficient injection pressure at greater depths [1]. Progressive Cavity Pumps (PCPs) are generally limited to approximately 6,000 ft since they are subject to torsional loads on the drive string and the differential pressure rating of the elastomeric stator [21, 22]. In the feature space, well depth provides good class separation: very deep wells are preferentially assigned to ESP or Gas Lift systems, while PCP installations cluster at shallower depths. Well depth is therefore a highly discriminative feature for the classifier.

3.4.2 Reservoir Pressure

Reservoir pressure constitutes the principal energy source driving fluids from the formation into the wellbore and establishes the differential pressure that the artificial lift system must provide. As reservoir pressure declines during depletion, the Inflow Performance Relationship (IPR) deteriorates such that greater lift energy is required to maintain the same production rate [1, 2]. Gas Lift systems are generally highly effective in high-pressure reservoirs, because the reduction in multiphase density is most efficient at substantial pressure drawdown [1, 2]. Conversely, in depleted or low-pressure reservoirs with low productivity indices, positive-displacement systems such as rod pumps and PCPs are preferable because they can reduce the flowing bottomhole pressure to near-zero values [13, 15]. ESPs can operate across a relatively broad pressure range, but they require sufficient pump intake pressure to prevent cavitation and gas-locking [20, 23].

A threshold behavior is also observed with Gas Lift efficiency: as reservoir pressure declines below the minimum pressure required to achieve critical flow through the operating valve, efficiency decreases rapidly. By contrast, rod pumps and PCPs maintain nearly constant volumetric efficiency even at low reservoir pressures. This nonlinear, method-dependent response to reservoir pressure constitutes a highly discriminating feature for the ML classifier, owing to the characteristic threshold behavior exhibited by Gas Lift at critical pressure and the comparatively stable low-pressure response of rod pumps and PCPs.

3.4.3 Production Rate

Production rate represents the volumetric throughput capacity of the lift system and is one of the most critical parameters governing artificial lift selection in practice. Modern ESP centrifugal stages can deliver rates ranging from approximately 200 to 30,000 STB/day under standard configurations, depending on pump series and casing diameter, depending on pump series and casing diameter [1, 14]. Rod pumps are better suited for low-to-moderate production rates below 5,000 STB/day (typically below 1,000 STB/day for most installations), plunger displacement volume, and the capital cost associated with larger surface units at higher outputs [2, 13] Gas Lift provides considerable operational flexibility, accommodating rates as low as 50 STB/day and as high as 50,000+ STB/day [1, 24]. PCPs are effective up to approximately 4,500 STB/day, with throughput limited by the maximum allowable rotational speed of the rotor–stator assembly [21, 22].

In the feature space, production rate establishes near-linear class boundaries with respect to lift method. High-rate wells are predominantly classified as ESP; low-rate wells are generally classified as rod pump; and intermediate-rate wells are distributed across the Gas Lift and PCP classes. This quasi-monotonic relationship yields high mutual information between this feature and the target label, thereby improving classifier performance.

3.4.4 Gas–Oil Ratio (GOR)

The gas–oil ratio (GOR) is a measure of how much gas is produced compared to oil per unit of stock-tank

oil, and it can significantly influence the performance of each artificial lift method. One of the main ways free gas affects artificial lift performance is at the intake of centrifugal pumps (ESPs), where its presence adversely impacts the head–capacity curve, induces surging, and in some cases causes gas locking when the in-situ gas void fraction exceeds approximately 10–15% at standard stage geometry [20, 23]. The presence of free gas also has a significant adverse effect on rod pump performance, as gas in the pump barrel reduces volumetric efficiency; operators therefore use gas anchors as a partial mitigation measure [2, 13]. Progressive Cavity Pumps (PCPs) generally tolerate moderate quantities of free gas more effectively than ESPs due to their positive-displacement operating principle; however, excessively high GOR reduces the effective displaced volume and accelerates elastomeric stator degradation through gas permeation [22, 25]. Conversely, Gas Lift exhibits the opposite response and generally performs more effectively at higher GOR because the produced formation gas can supplement the injected lift gas, reducing the required injection gas–liquid ratio and the associated compression costs [1, 24].

As Gas Lift efficiency improves with an increase in gas–oil ratio (GOR), ESP performance deteriorates due to the adverse effects of free gas on the head–capacity curve, including surging and potential gas locking, while rod pumping efficiency deteriorates due to reduced volumetric efficiency caused by free gas in the pump barrel; thus creating an excellent discriminative signal. The classes of Gas Lift are therefore accurately identified and separated from all pumped alternatives through this informative classifier, resulting in a high confidence level.

3.4.5 *Water Cut*

The degree of water contained in the total liquid production (water cut) influences the density of that liquid, as well as the likelihood of forming emulsions and the risk of corrosion - all these factors determine whether the deployed ALM is appropriate. Given the relatively low sensitivity of centrifugal pump performance to fluid specific gravity variations, ESPs are well-suited for high water cut conditions, provided sufficient flow rates are maintained for motor cooling. Rod pumps also tolerate high water cut, but increased rod string loading at elevated fluid densities promotes premature pump failure, and water-in-oil emulsions can significantly elevate effective fluid viscosity [2, 13]. Gas lift efficiency diminishes at very high-water cuts, as greater gas injection volumes are required to achieve the same hydrostatic pressure reduction in the denser water-dominated fluid [1, 24]. PCPs handle water-dominated fluids effectively but are susceptible to accelerated corrosion from produced water, particularly in the presence of CO_2 or H_2S [21, 22].

The nonlinear feature space boundaries arising from the interaction of water cut with other variables (viscosity, corrosion) are well-suited to tree-based and ensemble classifiers. Water cut enhances the model's ability to differentiate lift methods in mature, waterflood-dominated reservoirs.

3.4.6 *Fluid Viscosity*

Fluid viscosity represents the resistance of the produced fluid to flow and is one of the most consequential parameters governing AL selection. In PCPs, the positive-displacement rotor–stator mechanism experiences reduced volumetric slip at higher fluid viscosities, which improves pumping efficiency [21, 22, 29]. PCPs are particularly suitable for producing heavy oils with viscosities exceeding 1,000 cP, with documented applications up to 100,000 cP in some heavy oil developments. By contrast, rod pumps can only accommodate viscosities up to 200–500 cP; beyond this range, rod pump performance is adversely affected by inadequate travelling valve fillage during the upstroke and by significantly increased plunger friction In centrifugal pumps (ESPs), as viscosity rises above approximately two hundred centipoise (cP), performance degrades significantly due to elevated hydraulic friction losses within centrifugal stages, resulting in increased power consumption and higher motor temperatures associated with the greater mechanical energy required within the pump [19, 20]. Gas Lift systems are comparatively insensitive to fluid viscosity since they reduce fluid column density rather than displacing fluid mechanically; however, at higher viscosities, multiphase flow is retarded and the overall efficiency of the Gas Lift system is reduced [1, 24].

The Progressive Cavity Pump (PCP) class is almost exclusively discriminated at high viscosity values, ensuring that the classifier reliably identifies PCP-appropriate wells.

3.4.7 *API Gravity*

API gravity is an internationally accepted industry-standard index for the density of a petroleum liquid relative to water, and it is calculated as the inverse of specific gravity. Heavy oil (API $<$ 22 °API) is characterized by high viscosity and a

tendency toward emulsion formation and asphaltene precipitation, which commonly create operational challenges for several artificial lift methods. The use of Progressive Cavity Pumps (PCPs) is generally the preferred option for heavy-oil reservoirs, as PCPs handle viscous fluids effectively and can also tolerate the unconsolidated sand commonly encountered in heavy-oil environments [22, 26, 27].

While Electrical Submersible Pumps (ESPs) can be applied in heavy-oil wells, it is generally necessary to derate the pump curve to account for hydraulic friction losses, increased motor loading, and the elevated risk of overheating associated with emulsion formation [20, 27]. The installation of rod pumps in heavy-oil wells introduces additional operational challenges, including increased rod loading, reduced pump fillage, and impaired travelling valve performance [2, 13]. While Gas Lift can be applied across a broad range of API gravities, its effectiveness diminishes with very heavy crude oils, primarily because reduced gas solubility limits the extent of achievable fluid density reduction [1, 24].

Even though API gravity is physically correlated with viscosity, it still provides additional discriminative information because it also reflects susceptibility to asphaltene and paraffin deposition, which are failure mechanisms distinct from degradation caused purely by viscosity. Consequently, incorporating both features enriches the overall feature representation.

3.4.8 Well Deviation

Well deviation is the inclination of the wellbore from vertical that imposes distinct mechanical and operational constraints on artificial lift systems that depend on reciprocating or rotating rod strings. Rod pumps are adversely affected by high wellbore deviation because sucker rods bear against the low side of the tubing in deviated sections, causing accelerated abrasive wear, increased friction loads, and rod buckling during the downstroke. In horizontal wells, rod pump failure frequencies have been reported to increase by a factor of three to five relative to vertical wells [28, 29]. ESPs are comparatively well-suited for deviated and horizontal wells because the absence of a rod string eliminates rod-on-tubing wear, although the pump should be set in a section with inclination below approximately 70° to prevent excessive bearing side-loading [1, 20]. Gas lift is the most deviation-tolerant method because the mandrels and valves are static downhole components, and the injected gas–liquid mixture rises regardless of wellbore trajectory [1, 24]. PCPs are moderately tolerant of deviation but are susceptible to increased drive string wear at dogleg sections, which reduces run life [21, 22].

Well deviation introduces a near-categorical split in the feature space: highly deviated wells strongly favor Gas Lift and ESP, while near-vertical wells are compatible with all four methods. This behavior enhances decision-tree splitting efficiency in ensemble classifiers.

3.4.9 Bottomhole Temperature

Bottomhole temperature (BHT) affects the thermal integrity of essentially all major AL components. Typical ESP motors are rated around 250–300 °F; high-temperature variants are also available, but the motor winding insulation degrades exponentially as temperature increases [20, 30]. The produced fluid must provide sufficient convective cooling, which generally requires maintaining a minimum annular velocity of approximately 1 ft/s past the motor housing. Additionally, as the bottomhole temperature (BHT) increases, the available thermal margin is reduced [19, 20]. Among the four principal artificial lift systems, the Progressive Cavity Pump (PCP) imposes the most stringent thermal constraint, as both nitrile and fluoroelastomer stator compounds exhibit significant swelling and loss of mechanical properties at approximately 250–300 °F, and may debond from the housing under sustained high-temperature exposure [21, 22]. Rod pumps exhibit moderate thermal tolerance in their mechanical components; however, the elastomeric downhole pump barrel seals have been known to fail under high-temperature conditions [2, 13]. Gas Lift remains the preferred artificial lift method in high-BHT environments due to its metallic downhole components, which can tolerate elevated temperatures with reduced material degradation compared to other systems [1, 24].

Overall, the BHT is essentially a “hard” exclusion criterion — a strict screening threshold. The reason for this is that wells that exceed either the nitrile or fluoroelastomer stator temperature limit for the Progressive Cavity Pump (PCP) or the motor temperature rating for Electrical Submersible Pumps (ESPs) are automatically excluded from those lift classes. This type of threshold behavior fits very well with the use of rule-based splits in decision tree models using tree splits.

3.4.10 Sand Production

Sand production representing the co-production of formation sand grains together with reservoir fluids that results in erosion of downhole equipment caused by abrasive solids, plugging due to sand accumulation, and mechanical wear of equipment components when sand is produced with the same fluids through an Electrical Submersible Pump (ESP). When sand is produced using an ESP, it can cause wear on the centrifugal impellers and diffusers of the ESP, which leads to degradation of pump head and rapidly failing bearings due to vibration from the ESP. There have been examples in the literature where a measurable loss in head has been documented when the concentration of sand was greater than 50 parts per million (ppm) [20, 31, 32]. In rod pumps, sand accumulates above the plunger, causing the plunger to become stuck and accelerating wear of the valve and seat surfaces [2, 13]. Progressive Cavity Pumps (PCPs) exhibit inherent sand tolerance, as solid particles pass through the pump owing to the compliance of the elastomeric stator without catastrophic mechanical failure, making PCPs the preferred artificial lift option in sand-producing formations [22, 25]. Gas lift systems are highly sand-tolerant owing to the absence of moving downhole components; the injected gas is simply commingled with the sand-laden produced fluid stream [1, 24].

As a categorical binary feature, sand production effectively partitions the feature space: the presence of sand substantially reduces the probability of ESP and rod pump classification while increasing the likelihood of Gas Lift and PCP assignments, resulting in cleaner initial splits in tree-based classifiers.

3.4.11 H_2S Presence

Hydrogen sulfide (H_2S) promotes sulfide stress corrosion cracking (SSCC), drives hydrogen embrittlement, and accelerates pitting corrosion, so there are strict material-selection requirements that must be followed for essentially all AL equipment. High-strength carbon steel sucker rods are particularly susceptible to SSCC; the rods that are used must comply with the NACE MR0175/ISO 15156 hardness limits ($\leq$22 HRC), and even when the material is compliant, there is still a reduction in fatigue life under sour-service conditions [33, 34].

For ESP systems in H_2S environments, it is not sufficient to replace a single component; upgraded metallurgy is typically required, such as American Iron and Steel Institute (AISI) 316L stainless steel housings, together with corrosion-resistant stage coatings. Additionally, specially armored cable is generally required. Collectively, these requirements result in a significant increase in capital expenditure [20, 35]. For PCP systems, stators may respond adversely depending on the elastomer chemistry; they may swell or undergo chemical degradation after prolonged H_2S exposure, so fluoroelastomer or HNBR stator compounds are often the preferred option [21, 22]. Gas lift generally remains the most H_2S-tolerant option, partly because downhole valves can be manufactured from corrosion-resistant alloys, and also because there are no elastomers or electric motors present to constitute primary failure points [1, 24].

As a binary exclusionary feature, the presence of H_2S effectively penalizes AL methods with greater material susceptibility, while favoring Gas Lift. With this H_2S condition included, the model is less likely to recommend equipment configurations that are vulnerable to rapid corrosion-induced failure.

3.4.12 Formation Type

Formation type is a categorical variable that indicates the lithological and petrophysical conditions of the producing reservoir and therefore appears to determine fluid dynamics, completion methods, and artificial lift (AL) systems compatibility [1, 36]. Sandstone formations frequently contain an unconsolidated matrix, which means they represent a higher risk of sand production as well as an increased likelihood of heavy oil deposits; for these reasons, Progressive Cavity Pumps (PCPs) are preferred for use in sandstone reservoirs [25, 26]. By comparison, carbonate formations tend to be characterized by crude oils of higher API gravity, natural void space due to fractures and/or weakly cemented matrices, and variable gas–oil ratio (GOR) trends—all of which influence the relative suitability of Gas Lift and Electrical Submersible Pump (ESP) systems [15]. Additionally, shale formations are primarily developed via horizontal drilling and hydraulic fracturing, generally produce low-to-moderate rates of production, and initially have relatively high gas–oil ratios (GORs); therefore, depending on the level of depletion reached at that stage of production, either a Gas Lift or rod pump system will be the preferable operating environment [28].

Therefore, formation type serves to lend lithology-specific context to the classifier in ways that cannot be measured using continuous fluid

and reservoir characteristics, thus improving the distinction between classes; this is particularly important for classifying wells which exhibit similar operational characteristics regardless of their respective formation types and completion types.

3.4.13 *Exploratory Data Analysis*

The distribution of key input features across all artificial lift classes, as well as the inter-variable relationships among them, was evaluated through exploratory data analysis. Box plots were used to visualize the distributions of selected features, while correlation heatmaps were generated to quantify the degree of multicollinearity among the input features.

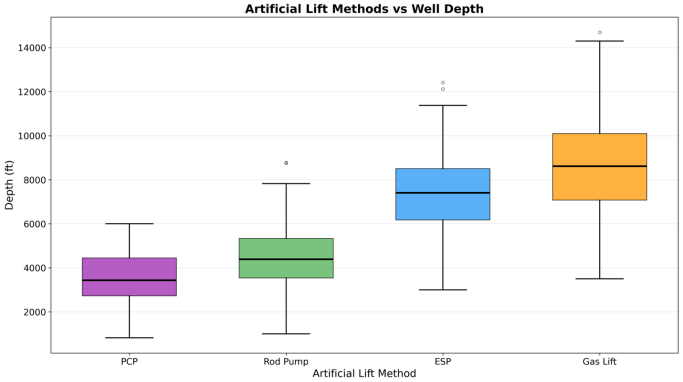

Figure 3. Box plot of artificial lift methods vs. depth (ft).

As illustrated in Figure 3, a distinct depth separation is observed among the four artificial lift classes. Gas Lift (median depth: approximately 8,381.99 ft) and ESP (median depth: approximately 7,606.59 ft) are the two artificial lift methods most frequently deployed at greater depths (i.e., above 5,000 ft). Conversely, Rod Pump (median depth: approximately 4,503.00 ft) and PCP (median depth: approximately 3,294.90 ft) are the two methods predominantly used in shallower completions (i.e., below 5,000 ft). These findings are consistent with established industry practice, wherein Gas Lift and ESP are frequently selected for deeper, higher-pressure wells, while Rod Pump and PCP are typically used in shallow-to-moderate depth wells.

As illustrated in Figure 4, a distinct association is observed between GOR and artificial lift class. Gas Lift wells exhibit the highest GOR values (median ≈ 863.06 scf/STB), whereas Rod Pump and PCP wells display comparatively lower GOR values (median ≈ 91.58 scf/STB and 43.30 scf/STB, respectively). ESP wells occupy an intermediate range (median ≈ 277.76 scf/STB), consistent with the broader operational envelope of centrifugal pump systems. Moreover, the low GOR observed for PCP wells is physically

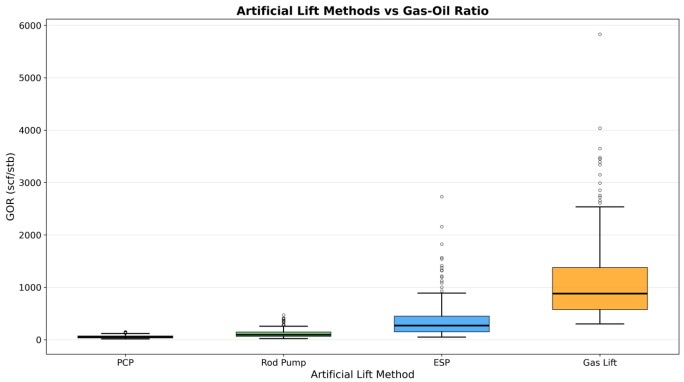

Figure 4. Box plot of artificial lift methods vs. GOR (scf/stb).

consistent with the predominant deployment of PCPs in heavy-oil reservoirs characterized by low associated gas content.

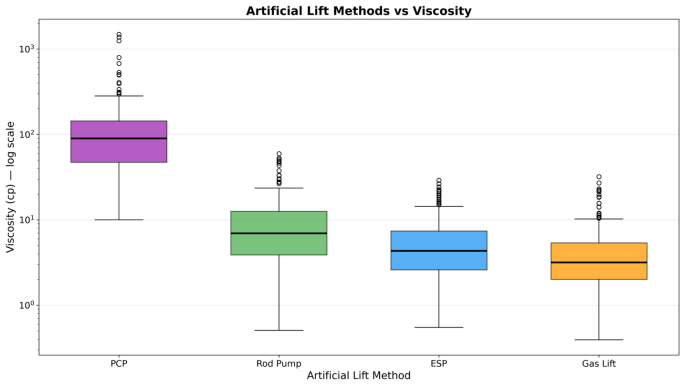

Figure 5. Box plot of artificial lift methods vs. viscosity (cp, log scale).

The relationship between artificial lift class and fluid viscosity is illustrated in Figure 5. It can be observed that PCP wells exhibit substantially higher viscosity than ESP and Gas Lift wells. The median viscosity of PCP wells is approximately 85.89 cP, compared to the median viscosities of ESP and Gas Lift wells, which are both below 10 cP (approximately 4.46 cP and 3.54 cP, respectively). Rod Pump wells exhibit a median viscosity of approximately 7.42 cP and therefore occupy an intermediate range relative to PCP and the other three artificial lift methods, reflecting the broader viscosity tolerance of reciprocating pump systems. These findings corroborate existing literature demonstrating the suitability of PCP systems for heavy, high-viscosity crude oils and the predominant application of ESP and Gas Lift to low-viscosity hydrocarbon production. A logarithmic scale is employed in the figure to clearly resolve the order-of-magnitude differences in viscosity across the four lift classes.

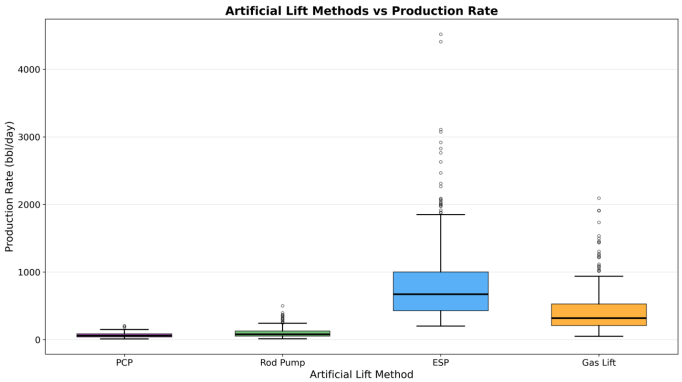

Figure 6. Box plot of artificial lift methods vs. production rate (bbl/day).

As shown in Figure 6, the median daily production rates of Rod Pump and PCP wells (approximately 96.85 STB/day and 56.08 STB/day, respectively) are substantially lower than those of Gas Lift and ESP wells (approximately 350.19 STB/day and 640.62 STB/day, respectively). This pattern is consistent with established industry practice, wherein ESP is the preferred choice for high-volume applications, Gas Lift for moderate- to high-rate wells, and Rod Pump or PCP for low-rate wells.

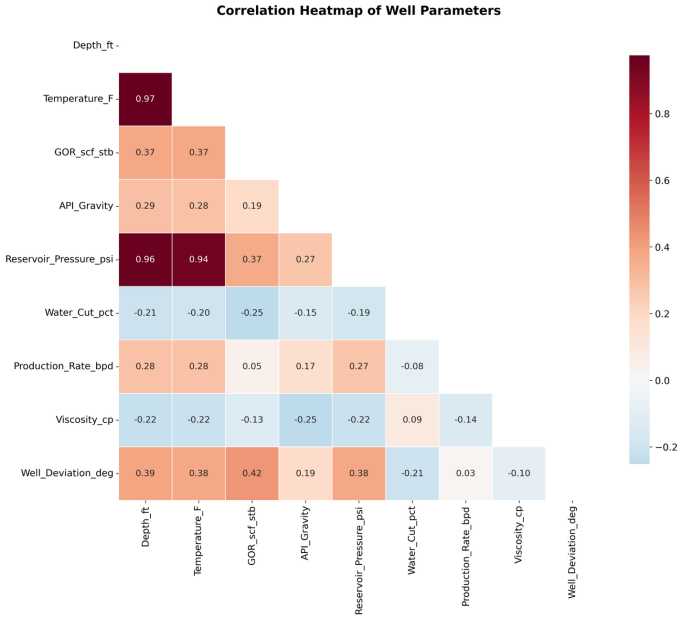

Figure 7. Correlation heatmap of well parameters.

Figure 7 presents a correlation matrix of the well input parameters. The relationships between depth and temperature (r = 0.97); reservoir pressure and temperature (r = 0.94); and depth and reservoir pressure (r = 0.96) exhibit strong positive correlations. These findings are consistent with the expected increase in hydrostatic pressure and the expected geothermal gradient with depth. A moderate positive association is also observed between production rate and each of these three parameters (depth: r = 0.28; temperature: r = 0.28; reservoir pressure: r = 0.17), whereas the correlation between production rate and GOR is comparatively weak (r = 0.05). Negative correlations can be found between viscosity and depth (r = -0.22), viscosity and temperature (r = -0.22), and viscosity and production rate (r = -0.14). A negative correlation can also be observed between water cut and depth (r = -0.21) and reservoir pressure (r = -0.19). Therefore, the results support retaining all twelve features used in model development as all of the input variables provide unique physical information to the feature space. PCA was used as a tool for exploration only to evaluate the intrinsic dimensionality of the training data and to visualize the internal structure within it [37].

3.5 Synthetic Data Generation and Class Balancing

Although The dataset for this project includes a total of 990 actual wells (400 ESP, 300 Gas Lift, 200 Rod Pump, and 90 PCP). However, a class imbalance existed in the training set as a result of the stratified train–test split (320 ESP, 240 Gas Lift, 160 Rod Pump, and 72 PCP). To address this imbalance, two independent oversampling techniques were used to augment the training data only. These included (1) SMOTE (Synthetic Minority Over-sampling Technique) and (2) DI-SDG (Domain-Informed Synthetic Data Generation). Both techniques were able to augment the underrepresented classes so that each of the four classes contained an equal number of samples for a total of 1,280 samples in the training set with balanced class representation. As a result, the test data remained the same because they were comprised of real well records only.

The DI-SDG process involves a series of six steps: (1) identify the target class which requires augmentation, (2) extract the parameter ranges for the corresponding artificial lift in Table 1, (3) choose an actual sample from the target class (this will serve as the seed point), (4) create perturbations by $\pm$10% of the seed point for each parameter within the valid parameter ranges, (5) determine if the candidate samples meet the screening criteria for all attributes, and (6) decide if the candidate samples are to be accepted or rejected. Generating synthetic data within these engineering and operational boundary conditions ensures that the resulting sample population will contain representative real-world samples [11, 12]. Synthetic samples violating any physical or operational

screening constraint were automatically rejected during augmentation.

3.5.1 Mathematical Formalization of DI-SDG

The DI-SDG algorithm is formalized through three core mathematical expressions that define the perturbation mechanism, distribution consistency condition, and hard constraint enforcement predicate.

Equation (1) defines the continuous feature perturbation mechanism, wherein each seed feature value is perturbed by a zero-mean Gaussian noise term and clipped to class-specific physical bounds:

$$\tilde{x}_j = \text{clip}\left(x_j^{\text{seed}} + \varepsilon_j, L_j^c, U_j^c\right) \tag{1}$$

where $\varepsilon_j \sim \mathcal{N}(0, \sigma_j^2)$.

In Equation (1), x_j^{seed} is the j-th feature of the seed sample drawn from class c, ε_j is a zero-mean Gaussian perturbation with standard deviation σ_j, and $[L_j^c, U_j^c]$ are the class-specific physical constraint boundaries derived from established artificial lift screening criteria (Table 1). The standard deviation σ_j is determined by one of two perturbation modes:

$$\sigma_j = \begin{cases} \alpha_j \cdot \max\left(|x_j^{\text{seed}}|, 1\right), & \text{Mode A (Relative)} \\ \beta_j \cdot (U_j^c - L_j^c), & \text{Mode B (Absolute)} \end{cases} \tag{2}$$

Mode A (Relative) applies $\sigma_j = \alpha_j \cdot \max\left(|x_j^{\text{seed}}|, 1\right)$ to features whose variability scales with magnitude (Depth, Temperature, GOR, API Gravity, Reservoir Pressure, Production Rate, Viscosity), where $\alpha_j \in [0.05, 0.25]$ is calibrated to published within-class variability estimates. Mode B (Absolute) applies $\sigma_j = \beta_j \cdot (U_j^c - L_j^c)$ to features with fixed physical bounds (Water Cut, Well Deviation). This replaces the original uniform $\pm 10\%$ rule with a physically justified, feature-specific specification; the complete set of scale parameters and their engineering justifications for all nine continuous features is provided in Table 2.

Equation (2) defines the distribution consistency condition, which formally requires that the augmented class-conditional distribution remains within the operationally validated domain:

$$\text{supp}(\tilde{X}_c) \subseteq \prod_{j=1}^{d} [L_j^c, U_j^c], \quad \forall c \in \Omega \tag{3}$$

Equation (2) is enforced through three complementary mechanisms: (i) hard-boundary enforcement via the clip operation in Equation (1) ensures no synthetic sample extends beyond known physical boundaries; (ii) seed-based induction preserves multivariate correlations inherent in observed training data, because correlated seed values are perturbed within narrow bounds (e.g., the geothermal coupling $T \approx 70 + 0.015D$ is automatically respected); and (iii) calibrated perturbation magnitudes matched to published within-class variability estimates ensure synthetic samples occupy the same region of feature space as real wells of the target class.

Equation (3) defines the hard constraint validity predicate that each synthetic sample must satisfy before acceptance into the augmented dataset:

$$\text{VALID}(\tilde{x}, c) \equiv \forall j \in \{1, \ldots, d\} : L_j^c \leq \tilde{x}_j \leq U_j^c \tag{4}$$

The class-specific bounds $[L_j^c, U_j^c]$ in Equation (3) are derived from published artificial lift screening criteria [20, 24]. Representative constraints include API Gravity 10–50°, Well Depth > 0 ft, Water Cut 0–100%, Reservoir Pressure > 0 psi, and Temperature > 60°F. If $\text{VALID}(\tilde{x}, c)$ fails after 10 perturbation attempts, the seed sample x_{seed} is inserted as a fallback, guaranteeing $|X_c| = N_t$ for all $c \in \Omega$ on every execution. This deterministic guarantee eliminates the need for post-generation rejection sampling.

(a). Justification of Perturbation Scale

A perturbation magnitude of $\pm 10\%$ ($\alpha = 0.10$) was selected based on both engineering rationale and empirical validation. Published within-class coefficients of variation for various petroleum reservoir parameters typically fall between 5% and 15% for well-characterized fields [24, 38]. A sensitivity analysis conducted across seven uniform α values ($\alpha \in \{0.03, 0.05, 0.08, 0.10, 0.15, 0.20, 0.25\}$) using a fixed Random Forest configuration confirmed that $\alpha = 0.10$ yields near-optimal macro-F_1 performance ($F_1 = 0.9391$, Accuracy = 0.9242). Furthermore, the feature-specific implementation outperformed all uniform α alternatives, demonstrating that the application of domain-calibrated perturbation scales yields superior classification performance.

The perturbation magnitude of $\pm 10\%$ was selected based on the expected amount of uncertainty and operational variations typically associated with field measurements in oil production systems. The uncertainty for API gravity values is typically ± 0.5°API, or API gravity values have $\pm 10\%$ variability due to sampling conditions for gas-to-oil ratio

Table 2. Feature-Specific Scale Parameters for DI-SDG.

Feature	Mode	Scale (α or β)	Justification
Depth_ft	Relative	$\alpha = 0.05$	Survey uncertainty $\sim 5\%$ [38]
Temperature_F	Relative	$\alpha = 0.05$	Geothermal gradient variation $\pm 5\%$
GOR_scf_stb	Relative	$\alpha = 0.10$	GOR varies $\sim 10\%$ within reservoir zone
API_Gravity	Relative	$\alpha = 0.05$	PVT measurement uncertainty ± 2–3°API
Reservoir_Pressure_psi	Relative	$\alpha = 0.08$	Pressure transient analysis uncertainty $\sim 8\%$
Water_Cut_pct	Absolute	$\beta = 0.10$	Range-based; bounded [0–98%]
Production_Rate_bpd	Relative	$\alpha = 0.15$	Rate varies significantly with choke settings
Viscosity_cp	Relative	$\alpha = 0.25$	Highly sensitive to temperature gradients
Well_Deviation_deg	Absolute	$\beta = 0.10$	Survey measurement uncertainty ± 2–3°

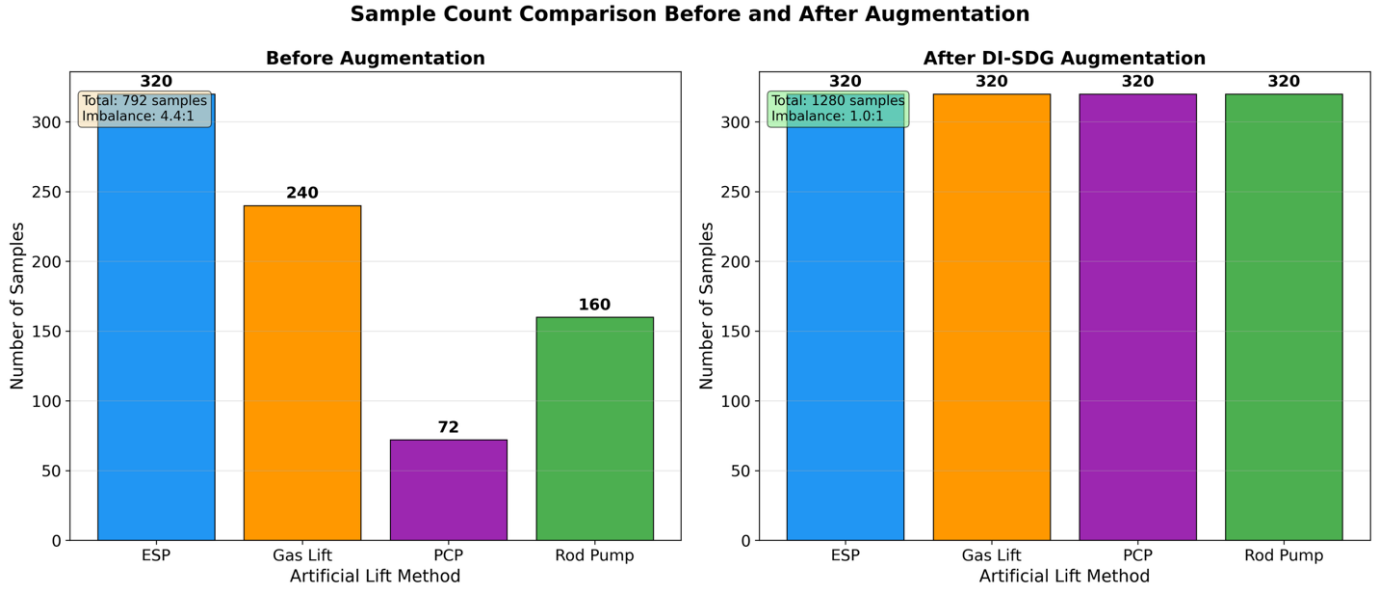

Figure 8. Sample count comparison in the training set before and after DI-SDG augmentation.

(GOR) and $\pm 2\%$ of full-scale accuracy for pressure gauges. This perturbation magnitude ensures that all synthetic samples retain their physical plausibility within physically plausible operating envelopes while introducing sufficient feature-space diversity between samples to allow generalization of the resulting models.

(b) Seed Sample Selection Strategy

The selected seed samples are selected via uniform random sampling with replacement from all available training samples of the target class c:

$$x_{\text{seed}} \sim u(x_c) \tag{5}$$

with random_state = 42 and was compared to a stratified alternative drawing seeds from quantile-based strata, resulting in a negligible performance difference ($\Delta F_1 < 0.003$), confirming the simpler random selection strategy is sufficient.

For categorical features (Sand Production and H_2S Presence), a Bernoulli flip is applied:

$$\tilde{x}_j = \begin{cases} 1 - x_j^{\text{seed}}, & \text{with probability } p_{\text{flip}} = 0.20 \\ x_j^{\text{seed}}, & \text{with probability } 0.80 \end{cases} \tag{6}$$

This formulation preserves dominant categorical class characteristics while introducing sufficient diversity.

Figure 8 illustrates the class distribution in the training set before and after DI-SDG augmentation, prior to augmentation (792 samples) with a class imbalance ratio of 4.4:1 comprising 320 ESP, 240 Gas Lift, 160 Rod Pump, and 72 PCP and after augmentation with 1,280 samples containing 320 samples each resulting in a balanced training set of 1,280 samples with a class balance ratio of 1.0:1 for each class. The DI-SDG method has successfully balanced the class distribution while preserving physically and operationally valid parameter combinations within established screening constraints [11, 12].

3.6 Feature Scaling and Dimensionality Reduction

Before developing the models, *StandardScaler* was used to standardize all numerical features. As an initial exploratory step, Principal Component Analysis (PCA) was employed. The purpose of PCA was to assess the degree of multicollinearity among correlated variables and visualize the structure of the training data in a reduced feature space. The model development and model evaluation were done using original variables instead of the principal components to maintain the interpretability of the individual features. Therefore, PCA is only to assess multicollinearity and characterize the overall structure of the training data and was not used as part of any feature selection or importance ranking process.

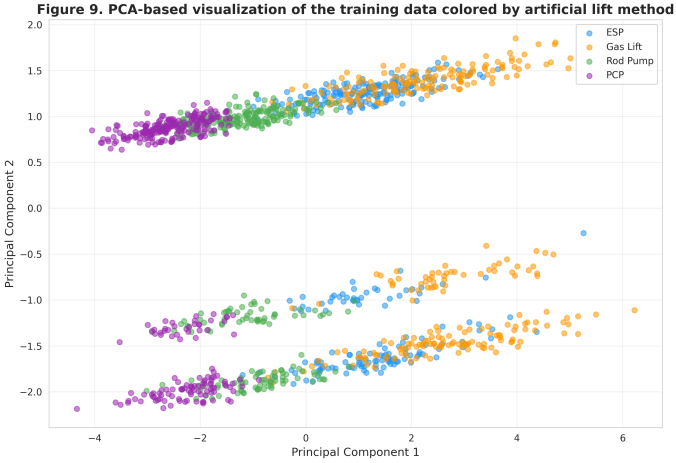

Figure 9. PCA-based visualization of the training data, colored by artificial lift class.

Based on the results that are presented in Figure 9 there is a degree of partial class separation among four classes of artificial lift in the reduced PCA space. PCP clusters appear relatively well-separated, while there is considerable overlap between ESPs and Gas Lift with Rod Pump occupying an intermediate region. The absence of complete linear separability in the projected feature space suggests the presence of nonlinear decision boundaries between the classes. This provides partial justification for the use of ensemble classifiers such as Random Forest, XGBoost, and Gradient Boosting, which can model the complex nonlinear class boundaries formed by combining the original input features.

3.7 Model Selection and Implementation

Six machine learning algorithms were selected because (1) they are effective in tabular classification tasks, (2) they support multi-class classification, and (3) they have precedent in engineering applications. A summary of the rationale for selecting each model is provided in Table 3.

3.8 Hyperparameter Optimization

Hyperparameter optimization was performed using GridSearchCV with 5-fold stratified cross-validation, with the objective function defined as macro-averaged F1-score [41]. The optimal hyperparameters identified for each model are presented in Table 4.

3.9 Evaluation Strategy

The analysis used five-fold stratified cross-validation for evaluating the models on the training set and a final evaluation on the test set (held out 20% of the original data, or 198 total samples). The evaluation included reporting of macro-averaged precision, recall, F1-score and overall accuracy. The use of macro-averaging was selected since it gives equal weight to all four artificial lift classes. This approach reduces bias in aggregate evaluation metrics due to the effects of majority-class dominance [42].

In addition to reporting on the various evaluation metrics for each of the models, confusion matrices were generated for the best performing model to identify specific misclassification patterns.

To perform the analysis, the original dataset was partitioned into an 80/20 stratified random split by class using stratify=y with the scikit-learn library's function train_test_split, setting a fixed random seed of 42. The benefit of stratified sampling is that the class distributions in both the training dataset (792 samples) and the testing dataset (198 samples) mirror the overall distribution of classes in the complete dataset (ESP: 40.4%, Gas Lift: 30.3%, Rod Pump: 20.2%, PCP: 9.1%) and prevents potential sampling artifacts that could cause any class to be over-represented or under-represented in either dataset. Given that there exists significant class imbalance in the dataset, an

Table 3. Rationale for selected machine learning models.

Model	Key Strengths	AL Selection Relevance
Decision Tree	Interpretable decision hierarchy; requires minimal preprocessing for mixed-type features	Mimics engineer-defined screening logic (e.g., depth and GOR thresholds)
Random Forest	Ensemble robustness; tolerates class imbalance; provides feature importance rankings	Quantifies the relative importance of parameters such as viscosity and GOR, supporting selection prioritization
K-Nearest Neighbors (KNN)	Instance-based; captures non-linear decision boundaries	Groups wells with similar operational profiles (e.g., depth–viscosity combinations)
XGBoost	Gradient boosting with regularization; state-of-the-art performance on tabular data	Demonstrated effectiveness in geoscience and petroleum engineering classification tasks [39]
Support VectorMachine (SVM)	Well-suited to high-dimensional feature spaces; robust to outliers	Resolves complex class boundaries in multi-parameter lift screening
Gradient Boosting	Sequential error correction; high accuracy on structured data	Captures subtle parameter interactions critical to lift method discrimination [40]

Table 4. Optimized hyperparameters for machine learning models.

Model	Key Hyperparameters	Optimal Values
Decision Tree	Criterion, Max Depth, Max Features, Min Samples Leaf, Min Samples Split	Entropy, None, None, 1, 2
Random Forest	n_estimators, Max Depth, Max Features, Min Samples Split	200, None, sqrt, 2
KNN	n_neighbors, Weights, Metric	7, Distance, Manhattan
XGBoost	Learning Rate, Max Depth, n_estimators	0.3, 3, 200
SVM	C, Kernel, Gamma	10, RBF, Scale
Gradient Boosting	Learning Rate, Max Depth, n_estimators	0.1, 5, 200

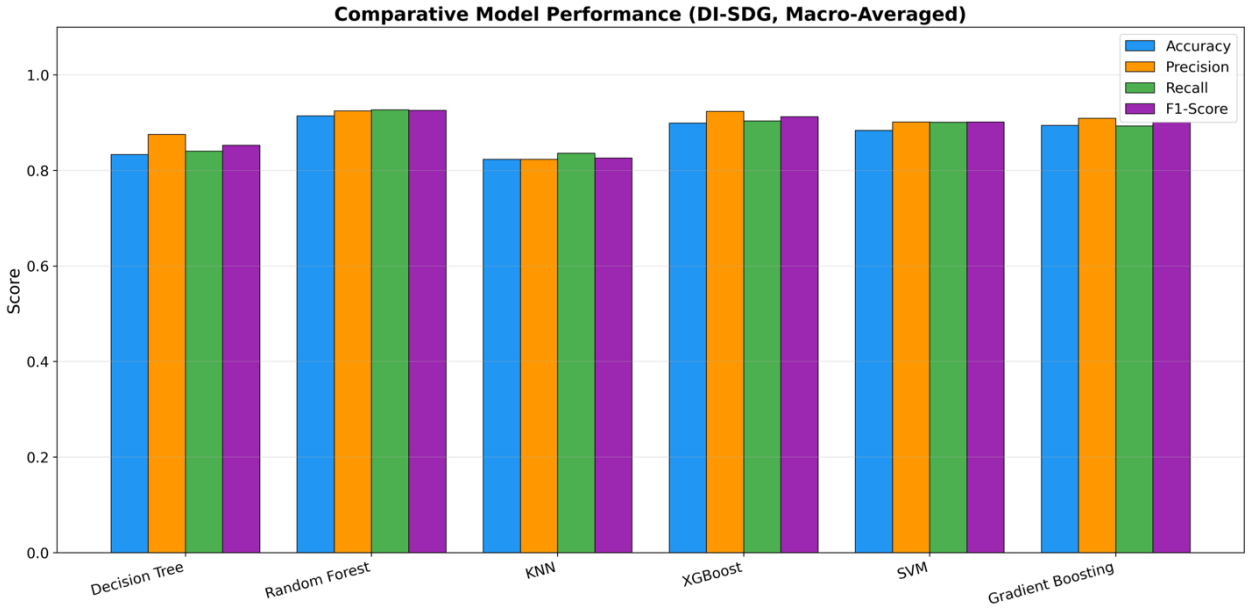

Figure 10. Comparative model performance (macro-averaged metrics, DI-SDG).

unstratified split would have resulted in the small number of PCP samples (minority class) possibly being insufficiently represented in the test set, thus compromising the reliability of evaluation metrics for

the PCP class.

Two additional validation strategies are acknowledged as providing stronger evidence of generalization beyond the present stratified random split. Time-based splitting — training on data from earlier wells and testing on data from later wells — would enable an evaluation of temporal generalization and robustness to concept drift as reservoir conditions evolve with depletion (e.g., progressive increases in water cut and declines in reservoir pressure). This approach simulates the realistic operational scenario in which a model trained on historical production records is applied to guide completion or intervention decisions on newly drilled or recently completed wells. Field-based splitting — withholding all wells from one or more complete fields as the test set — constitutes the most rigorous validation framework for field deployment, as it simulates the application of a model trained on existing fields to a previously unseen field with potentially distinct geological characteristics. Both strategies require timestamped or field-labelled production data, which are not available in the present synthetic dataset. The acquisition of real operator data with temporal and spatial metadata to enable these validation approaches is identified as the primary direction for future research.

4 Results and Discussion

4.1 Comparative Model Performance

Table 5 presents the macro-averaged evaluation metrics for all six models trained on the DI-SDG-augmented dataset and evaluated on the held-out test set. Figure 10 provides a corresponding visual comparison of these metrics across all classifiers.

Table 5. Comparative evaluation metrics (macro-averaged) for six ML models on DI-SDG augmented data.

Model	Accuracy	Precision	Recall	F1-Score
Decision Tree	0.8333	0.8752	0.8403	0.8528
Random Forest	0.9141	0.9247	0.9267	0.9255
KNN	0.8232	0.8234	0.8358	0.8258
XGBoost	0.8990	0.9234	0.9035	0.9125
SVM	0.8838	0.9015	0.9007	0.9011
Gradient Boosting	0.8939	0.9089	0.8931	0.9006

Random Forest achieved the highest performance across all evaluated metrics, recording an accuracy of 0.9141, a precision of 0.9247, a recall of 0.9267, and a macro-averaged F1-score of 0.9255. XGBoost and Gradient Boosting followed with F1-scores of 0.9125 and 0.9006, respectively. SVM produced intermediate results with a macro-averaged F1-score of 0.9011 (90.11%), while Decision Tree and KNN yielded the lowest macro-averaged F1-scores of 0.8528 (85.28%) and 0.8258 (82.58%), respectively. These results indicate that Random Forest and other ensemble classifiers provide a measurable advantage over individual base models in predicting the most suitable artificial lift method for a given set of reservoir, fluid, and operating conditions — an outcome consistent with the complex nonlinear interactions that characterize the relationships among these variables.

In the case of KNN classification, the performance was relatively low, indicating that distance-based classification may not be as effective in heterogeneous, multi-parameter spaces even after preprocessing and class balancing have been completed. The feature importance analysis presented in Figure 11 corroborates the Random Forest results, identifying GOR, viscosity, production rate, well deviation, and depth as the most influential predictors. These variables are consistent with long-established artificial lift screening criteria, confirming that the model successfully captured physically meaningful drivers of lift-method selection [1, 2].

4.2 Impact of Synthetic Data Generation: SMOTE vs. DI-SDG

The performance of the proposed Domain-Informed Synthetic Data Generation method was evaluated through a controlled comparative study against SMOTE. Each method was applied exclusively to the imbalanced training set to generate the requisite number of synthetic samples for each of the four artificial lift classes (320 samples per class), yielding a balanced training set of 1,280 samples. The test set remained unmodified and contained solely authentic well records; all six models were trained and evaluated under identical experimental conditions.

As previously referenced, $k\_neighbors = 5$ and sampling_strategy = 'not majority' are the SMOTE default parameters. Therefore, both $k\_neighbors$ and sampling_strategy were established for reproducibility using random_state = 42. By conducting sensitivity analysis with varying values from $k \in \{3, 5, 7, 9, 11\}$, it was possible to verify that $k = 5$ yields the highest macro-F_1 score among all tested values. Therefore, SMOTE and DI-SDG were applied using the same hyperparameter grid in a manner that ensured experimental parity across both augmentation conditions.

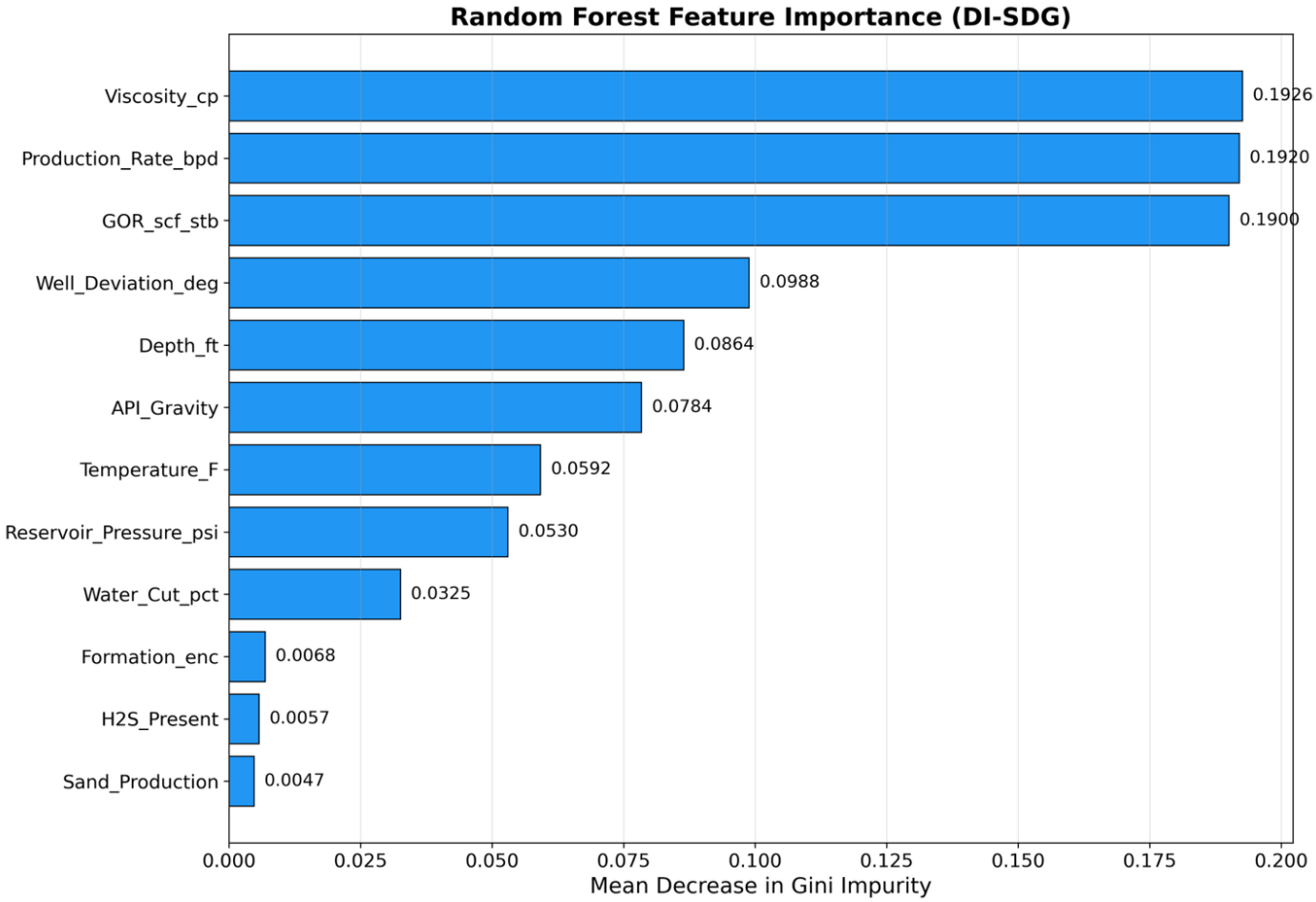

Figure 11. Random Forest feature importance for artificial lift selection.

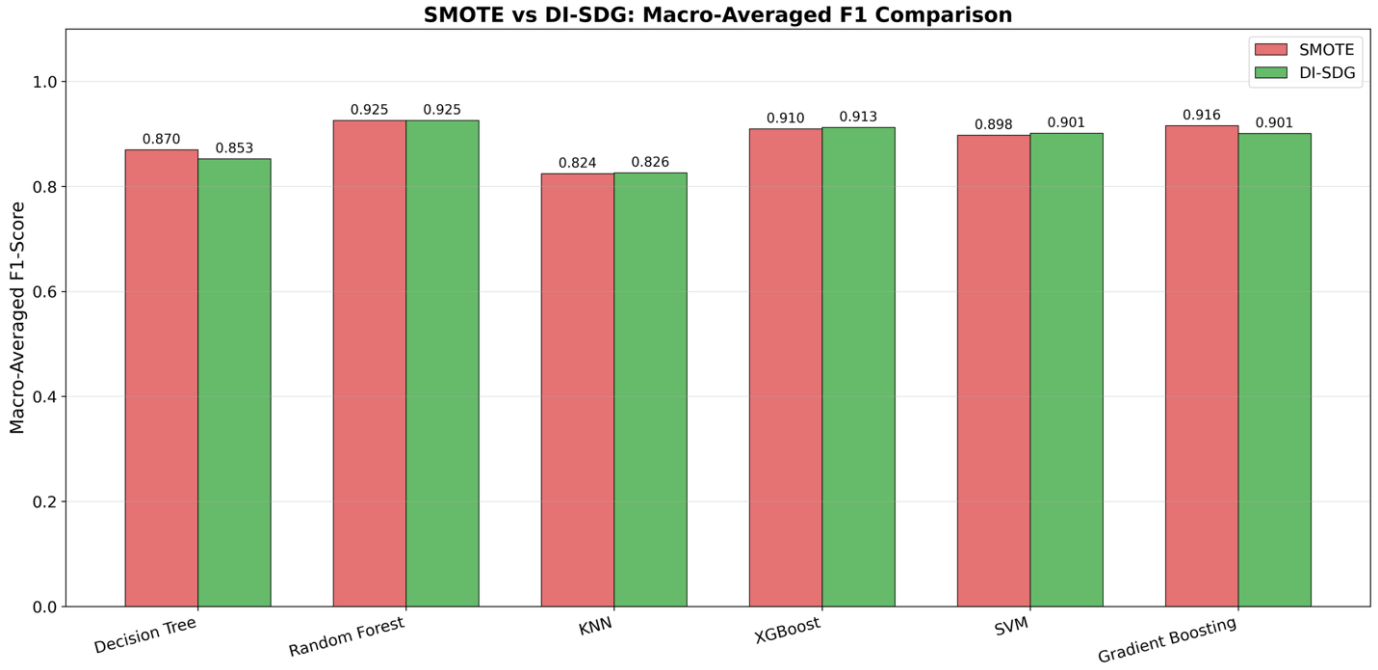

Figure 12. SMOTE vs. DI-SDG comparison (macro-averaged F1-score).

The results displayed in Table 6 and Figure 12 higher for DI-SDG than SMOTE for KNN, XGBoost, demonstrate that macro-averaged F1-scores were and SVM. There was parity between DI-SDG

Table 6. Comparison of macro-averaged F1-scores for SMOTE and Domain-Informed Synthetic Data Generation (DI-SDG).

Model	SMOTE F1	DI-SDG F1	F1 Δ (DI-SDG – SMOTE)
Decision Tree	0.8697	0.8528	−0.0168
Random Forest	0.9255	0.9255	0.0000
KNN	0.8243	0.8258	+0.0014
XGBoost	0.9099	0.9125	+0.0027
SVM	0.8976	0.9011	+0.0035
Gradient Boosting	0.9157	0.9006	−0.0151

and SMOTE for Random Forest (with respect to macro-averaged F1-score), and there were marginally lower F1-scores for Decision Tree and Gradient Boosting with DI-SDG as compared with SMOTE. Since DI-SDG does not uniformly outperform SMOTE across all classifiers but has also demonstrated a high level of competitiveness with SMOTE; additionally, DI-SDG is the only augmentation option that offers a physically constrained augmentation alternative to purely statistical oversampling.

In summary, DI-SDG provides a significant advantage due to all the samples produced under the DI-SDG process being bounded within established artificial lift screening criteria (thereby producing synthetic samples that are physically and operationally plausible); while SMOTE relies on linear interpolation within the feature space; subsequently generating synthetic samples that are mathematically valid, but may violate physical engineering constraints. This distinction can have major implications in petroleum engineering applications, as synthetic training data must reflect valid physical relationships among measured well and reservoir parameters in order for a resultant model to be reliable in actual applications [11, 12].

When combined with Random Forest, DI-SDG demonstrated its strongest relative advantage compared to any other classifier used in this study. This indicates that domain-constrained augmentation is most valuable when the classifier takes into consideration the nonlinear relationships between physical parameters. Therefore, DI-SDG is not considered to be an overarching replacement for SMOTE, but rather a domain-knowledge-based augmentation technique that proves particularly effective when used in conjunction with ensemble classifiers operating on physically structured feature spaces. DI-SDG effectively balances statistical performance and engineering validity.

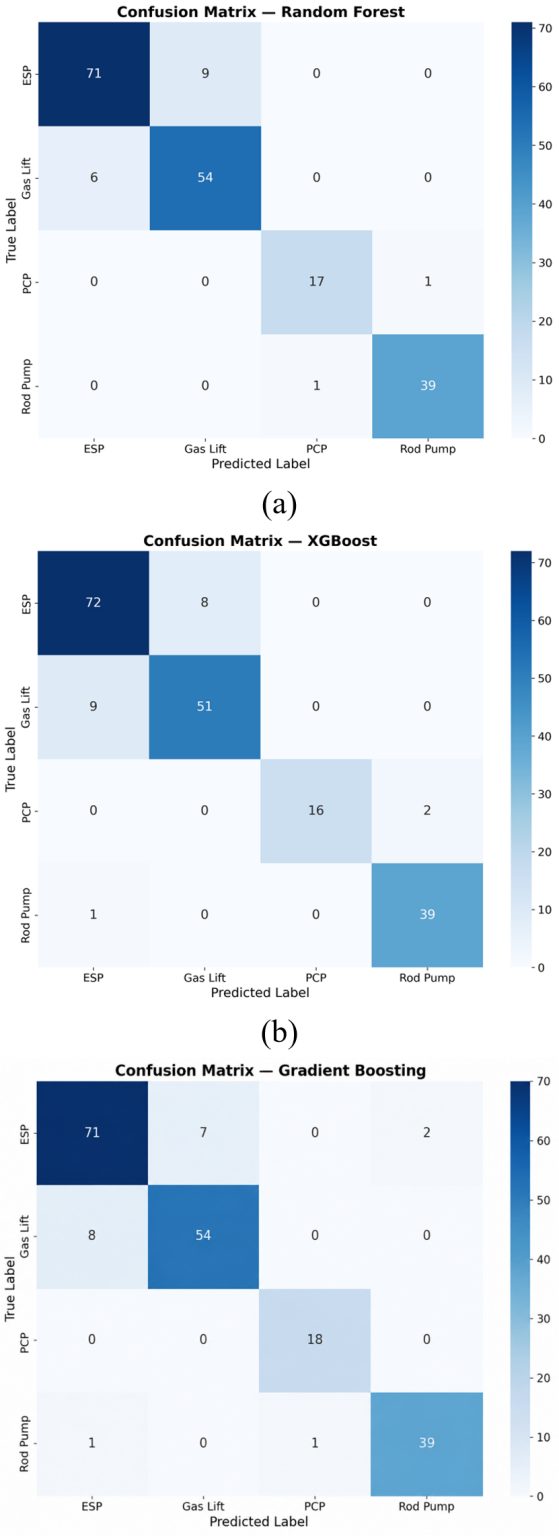

Figure 13. Confusion matrices for the top three performing models trained on DI-SDG-augmented data. (a) Random Forest; (b) XGBoost; (c) Gradient Boosting.

4.3 Confusion Matrix Analysis

The confusion matrices presented in Figure 13 enable assessment of the classification performance of the three top-performing models — Random

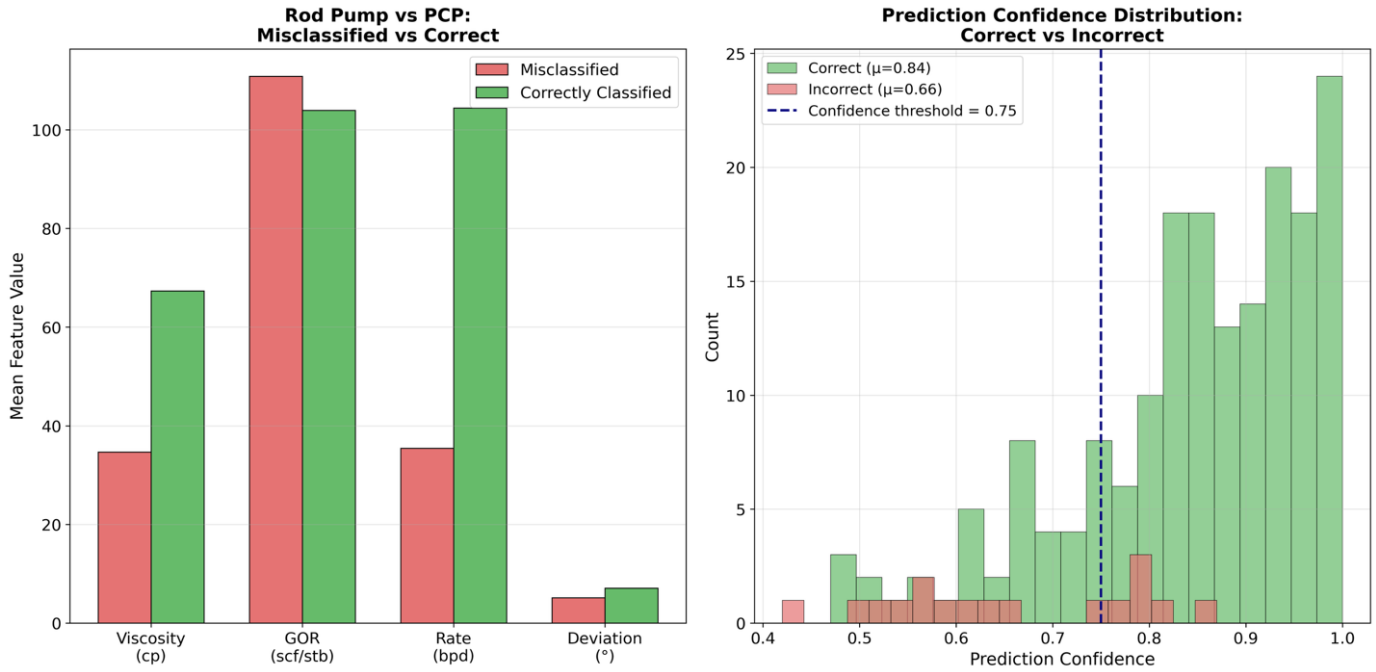

Figure 14. Misclassification pattern analysis for the Random Forest model: (a) mean feature values for misclassified vs. correctly classified PCP and Rod Pump wells; (b) prediction confidence distributions for correct and incorrect classifications, with a triage threshold at 0.75.

Forest, XGBoost, and Gradient Boosting — trained on DI-SDG-augmented data, at the individual well level. Examination of these matrices identifies which artificial lift classes contribute most to misclassification errors for each model.

The classifiers produced high classification accuracy for ESP and Gas Lift and much lower classification accuracy for Rod Pump and PCP. This is likely due to the distinct operational signatures of each lift class; ESP systems have distinct characteristics as they are typically found on high-rate wells that operate at moderate GOR, and Gas Lift has unique characteristics since it is primarily used on high-GOR environments. In evaluating the performance of different lift types, it is critical to consider overlap in their classifications across the three classifiers among Rod Pump and PCP wells as shown in Figure 14. The classification overlap is consistent with the engineering practice of using both types of pumps interchangeably in shallow to moderate-depth reservoirs where sand production rates and fluid viscosity are two key parameters that may occupy similar ranges for borderline wells. The Random Forest classifier demonstrated greater robustness in handling ambiguous boundary cases than the other classifiers; however, residual classification uncertainty persisted among all four lift classes.

A systematic misclassification analysis was conducted on all 18 prediction errors. The dominant confusion zone involved ESP and Gas Lift (15 cases: 8 Gas Lift wells misclassified as ESP and 7 ESP wells misclassified as Gas Lift), where misclassified wells exhibited a mean GOR of approximately 720 scf/stb (range: 480–1,420 scf/stb) and mean well deviation of 32°, placing all 15 cases within a physically overlapping region where both methods are viable and final selection is governed by surface infrastructure availability, capital expenditure constraints, and operator preference — factors not encodable in the 12-feature input space. The model's uncertainty in this zone therefore reflects genuine engineering ambiguity rather than a model deficiency. The secondary confusion involved PCP wells misclassified as Rod Pump (2 cases), where both misclassified wells exhibited intermediate viscosity (mean ˜87 cP, compared to ˜420 cP for correctly classified PCP wells and ˜35 cP for correctly classified Rod Pump wells), absent sand production, and shallow depth below 3,500 ft; the absence of a positive sand production flag removes the primary discriminating signal between these two methods. One additional error involved a Rod Pump well misclassified as ESP, representing an isolated boundary case. Critically, zero confusion was observed between ESP and PCP or between Gas Lift and PCP across all 198 test samples, confirming that the model reliably identifies extreme cases and concentrates uncertainty exclusively within physically

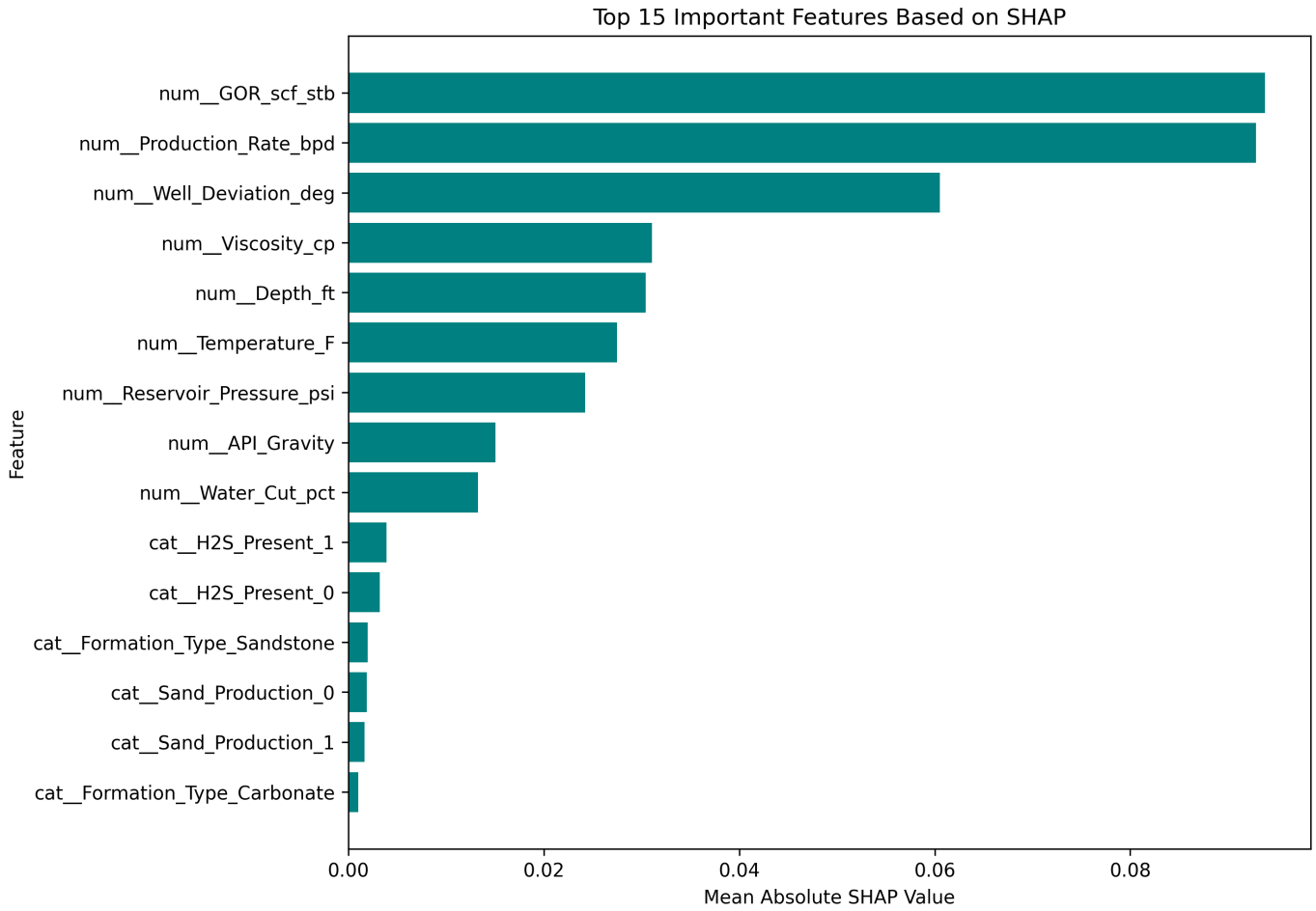

Figure 15. Mean absolute SHAP feature importance for the Random Forest model, showing the top predictors of artificial lift method selection.

meaningful transition zones. For practical deployment, three field guidance rules are recommended: (a) ESP or Gas Lift predictions with confidence below 0.75 and GOR between 400–1,800 scf/stb should be referred for detailed techno-economic analysis; (b) Rod Pump or PCP predictions with confidence below 0.75 on shallow, moderately viscous wells should prioritize sand production history as the primary override criterion; and (c) a confidence threshold of 0.75 is recommended as a general triage tool, supported by a statistically significant 18 percentage point gap between mean confidence for correct classifications ($\mu = 0.82$) and misclassifications ($\mu = 0.64$) seen in Figure 14.

4.4 Model Interpretability and Feature Importance

To enhance the transparency of the Random Forest model as a practical decision support tool, SHAP (SHapley Additive exPlanations) analysis was performed on the best-performing Random Forest model to improve interpretability. As illustrated in Figure 15, the mean absolute SHAP values indicate that GOR exhibits the highest mean absolute SHAP value and is therefore the most influential predictor in the Random Forest model, followed in order by production rate, well deviation angle, viscosity, and depth. In accordance with established artificial lift screening principles, the results are physically consistent in that GOR serves as a primary gas lift suitability indicator; production rate governs the applicability of ESP systems; and both deviation angle and viscosity reflect the mechanical constraints relevant to Rod Pump systems [1, 2].

Figure 16 presents the SHAP summary plot, which extends the bar chart analysis by showing the direction and distribution of each feature's contribution across all individual test samples. Each point represents one well, with position along the x-axis indicating the magnitude and direction of the SHAP value (positive values push the prediction toward a given class; negative values push away), and color encoding the original feature value (high values in one color, low in another). The summary plot confirms that high GOR

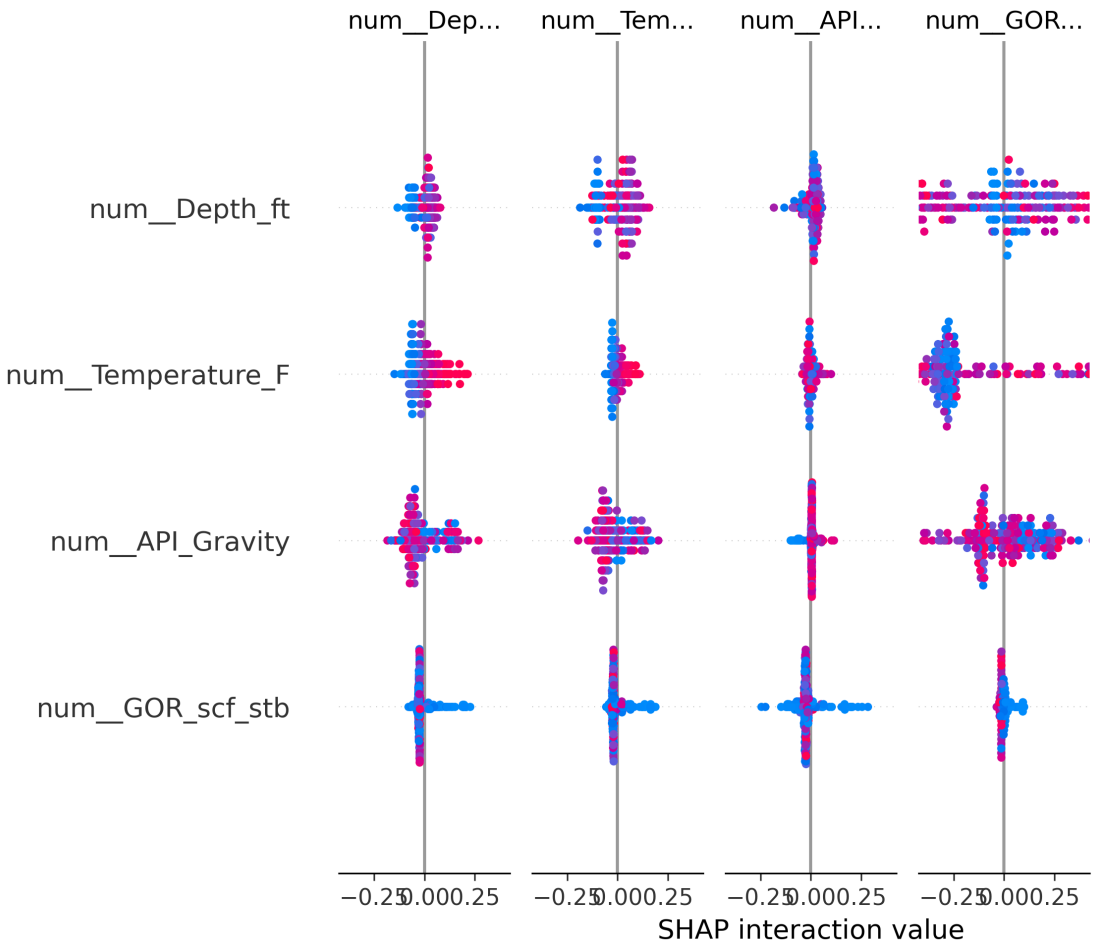

Figure 16. SHAP summary plot.

values consistently increase the probability of Gas Lift classification while suppressing ESP and Rod Pump assignments. High production rate values strongly favor ESP classification. High viscosity values are the dominant signal for PCP identification, with near-zero SHAP contributions to all other classes. Well deviation angle shows a bifurcated pattern: high deviation increases Gas Lift and ESP probability while markedly reducing Rod Pump probability, consistent with the known mechanical limitations of reciprocating rod strings in deviated wellbores. Formation type and H_2S presence exhibit comparatively diffuse distributions, corroborating their lower mean absolute SHAP values and their role as secondary modifiers rather than primary discriminators in the classification task [1, 2].

4.4.1 *H_2S Presence*

The prevalence of H_2S presence, as shown in Figure 15, varies across the four different artificial lift method groups, indicating that the presence of H_2S is 15% in ESPs and 41% in PCPs. However, the H_2S feature has a mean absolute SHAP value of 0.0039, indicating it has limited discriminative power among the four artificial lift methods, and ranks 10th among the 12 input features in terms of importance. This apparent paradox between statistical association and low predictive importance is consistent with H_2S's role in artificial lift engineering. H_2S is a parameter that influences the specification of the selected lift method; it does not fundamentally constrain the choice of method itself.

A statistical analysis using the chi-square test for independence has been performed, demonstrating that the presence of H_2S has a statistically significant association with the type of artificial lift ($p < 0.001$, $df = 3$). However, the effect size is small, which is consistent with the near-uniform prevalence rates of H_2S across the four artificial lift methods (15%–41%). In addition, a clear majority of the data does not contain the H_2S feature, with only 22% of all wells being classified as H_2S-present and 78% classified as H_2S-absent. Because of the near uniformity of distribution of the H_2S feature across artificial lift methods, it is concluded that the presence of H_2S provides limited discriminative power for lift method classification.

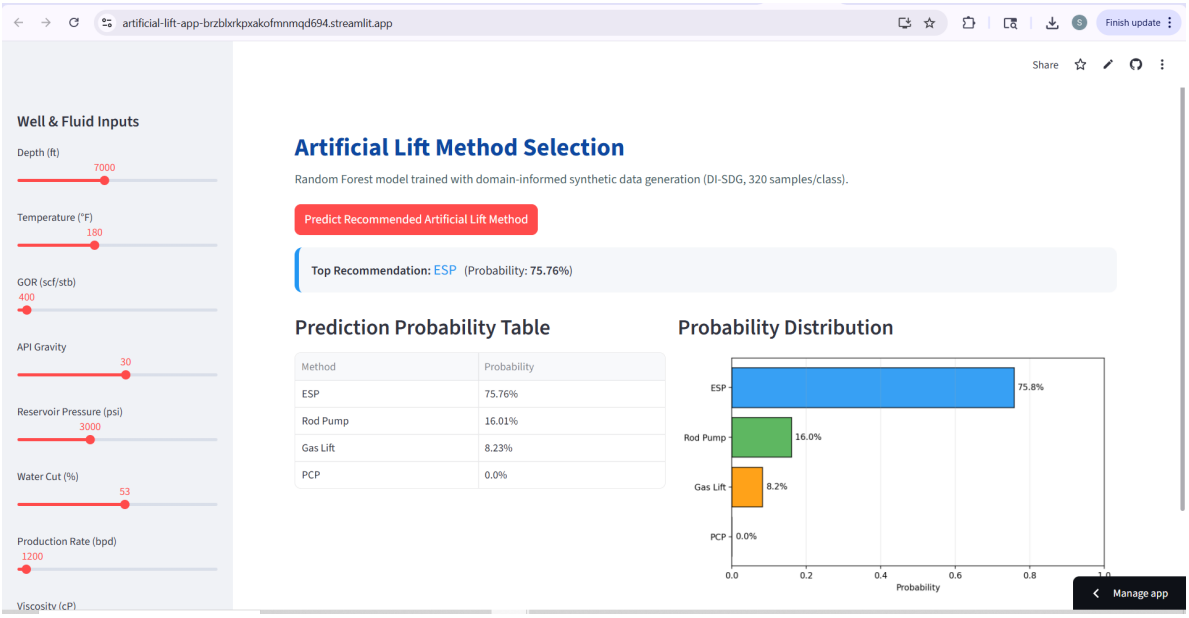

Figure 17. Streamlit-based decision-support interface developed for artificial lift method selection, showing user inputs, predicted lift recommendation, probability table, and class probability distribution.

In sour-service environments, there are industry requirements for material selection to prevent sulfide stress cracking and hydrogen embrittlement NACE MR0175/ISO 15156 [2, 33]. There are metallurgical specifications for each of the four principal lift technologies to meet these requirements. Corrosion-resistant alloys, specialized coatings, and enhanced protector designs, enable rod pumps, gas lift valves, and PCPs to operate successfully in H_2S-bearing wells through analogous material modifications [43, 44]. Consequently, H_2S presence functions as a specification modifier rather than a method discriminator within the lift selection decision framework [1].

This is consistent with the engineering reality that H_2S presence primarily governs material grade selection — for example, the requirement for NACE MR0175 [33]-compliant metallurgy — rather than the choice of lift mechanism itself, which explains the divergence between algorithmic feature importance and engineering significance. It is important to emphasize that this finding pertains exclusively to lift method selection; H_2S remains a critical consideration in equipment design, materials procurement, safety classification, and operational cost estimation following method selection [2]. The incorporation of sour-service cost differentials — such as the premium associated with NACE MR0175 [33]-compliant components — into an enhanced techno-economic selection framework is identified as a valuable direction for future research, as it would reintroduce H_2S as a cost-weighted decision variable with an increased predictive contribution.

4.5 Prototype Decision-Support Application

A prototype decision-support application was developed to demonstrate the practical utility of the machine learning framework described in this study. The application was built using Streamlit and deployed around the best-performing Random Forest model trained on the DI-SDG-balanced dataset. Users can enter twelve well and fluid parameters — wellbore depth, reservoir temperature, GOR, API gravity, reservoir pressure, water cut, production rate, viscosity, sand production, well deviation angle, H_2S presence, and formation type — and receive a prediction of the most suitable artificial lift method for the given conditions. The application also displays class probabilities for each candidate lift method in both tabular and graphical formats, allowing the practicing engineer to interpret results at a glance.

The tool was validated against two blind test wells (BT-1 and BT-2) withheld entirely from model training, achieving correct predictions in both cases. These results provide initial confirmation of the application's reliability ahead of wider field deployment (Figure 17).

4.6 Blind Well Test Validation

The Random Forest model's ability to generalize in real-world situations was evaluated by subjecting it to two independent field cases from the peer-reviewed

Table 7. Blind test well dataset: independent field cases from the published literature used for external validation of the Random Forest artificial lift classifier.

Parameter	Blind Test Well-1 Well A [45]	Blind Test Well-2 Yaran Field [46]
Well Depth (ft)	12,440	12,000
Reservoir Pressure (psi)	4,900	3,500
Production Rate (STB/day)	3,325	5,600[a]
GOR (scf/STB)	559	500
Water Cut (%)	92	30
Viscosity (cP)	0.55	2.00*
API Gravity (° API)	29.6	39
Well Deviation (°)	65*	5*
Bottom-Hole Temp (° F)	270	200
Sand Production	No*	No*
H_2S Presence	No*	No*
Formation Type	Sandstone (Nubia Fm.)	Carbonate (Sarvak/Fahlian Fm.)
Documented AL Method	Gas Lift	ESP
Model Prediction	Gas Lift – 61.3%	ESP – 58.75%

*Assumed or derived values; see Appendix A. [a] Assumed value derived from reported in see Appendix A.3.

literature. These cases were entirely independent of every phase of the model's development, including dataset construction, training, hyperparameter tuning, and hold-out evaluation. The main research question addressed with these two cases was whether the Random Forest model would be able to predict the artificial lift method used in wells it had never seen previously, using only the twelve static variables available at the time of lift selection.

The two cases used for the test were specifically selected to provide a diverse array of conditions. The first case was a mature Gas Lift well with a high water cut producing from the Nubian Formation in the Gulf of Suez, as described by Shedid and Yakoot [45].The second case was an installed ESP in the Yaran Field, which is a deep carbonate reservoir located in southwestern Iran [46]. The broad range of reservoir types and lift methods covered by the two cases, as well as their geographic separation, represents a significant challenge for the model's ability to generalize previously learned decision-making patterns to unseen reservoir conditions.

The blind test well parameters are compiled in Table 7. For several inputs not explicitly reported in the source publications, values were derived from published correlations, regional analogues, or established engineering conventions; these entries are marked with an asterisk (*). Full derivations are provided in Appendix A. The assembled feature vectors are considered sufficiently representative of the documented field conditions to support a meaningful external validation.

4.7 Engineering Analysis of Blind Test Cases

The documented lift selections for both wells can be seen to represent the screening criteria and the operating envelope parameters discussed in Section 3.4. Therefore, based on the feature space used to create this trained machine-learning model, it has adequately encapsulated the actual engineering conditions of these blind test cases.

Well A [45] is a mature offshore producer with a very high water cut (92%) and a moderate gas-oil ratio (GOR) of 559 scf/STB. This combination suggests a lift method capable of handling large fluid volumes with minimal sensitivity to free gas. A recognized advantage of Gas Lift over Electric Submersible Pumps (ESP) is its ability to operate effectively in gas-prone, high-water-cut environments, and the offshore setting with elevated bottomhole temperature (270°F) further favors a surface-controlled system with limited thermally sensitive downhole components. The documented injection depth of 4,570 ft is consistent with standard Gas Lift mandrel design practice, providing adequate pressure drawdown without excessive compression requirements [24].

The Yaran Field well [46] presents a different set of conditions: a deep onshore carbonate reservoir

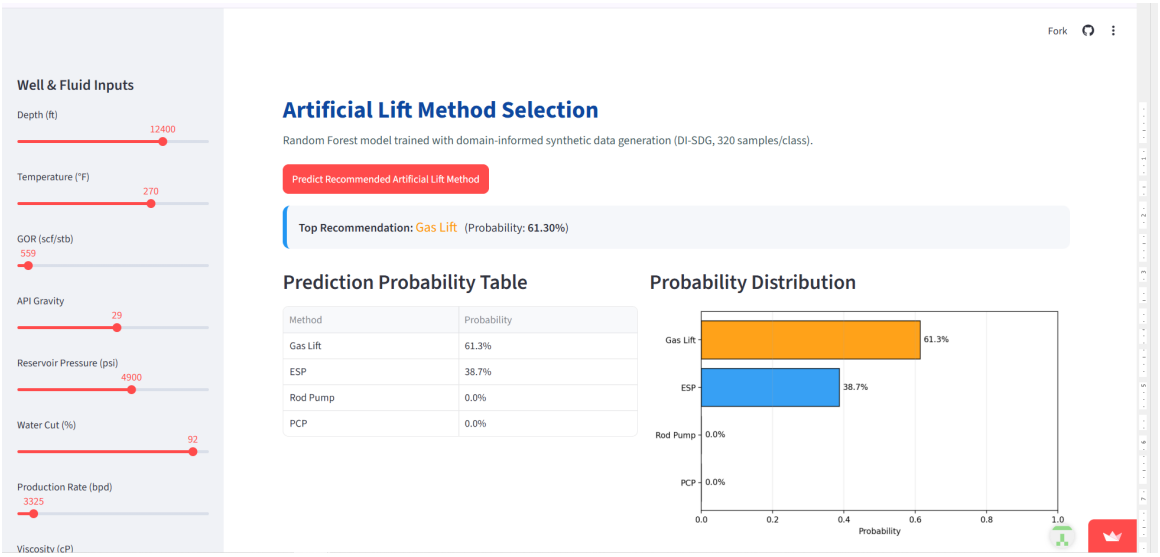

Figure 18. Well A Streamlit validation.

(∼12,000 ft), low water cut (30%), low GOR (500 scf/STB), moderately light crude (39°API), and estimated low viscosity (2 cP). The estimated production rate of 5,600 STB/day, derived from the reported Productivity Index of 2 STB/day/psi at an aggressive drawdown typical of ESP installations (refer AppendixA.3), falls well within the high-rate throughput range for which ESP is the preferred lift method [20]. The near-vertical trajectory, absence of sand, absence of H_2S, and consolidated carbonate formation present no mechanical or metallurgical constraints on ESP deployment. Taken together, these characteristics place this well firmly within the ESP operating envelope.

4.7.1 Model Prediction Results and Discussion

The trained Random Forest classifier was run through the Streamlit application using the feature vectors in Table 7. Prediction outputs and full class probability distributions for both wells are shown in Figures 18 and 19.

For Blind test well-1; Well A, the model predicted Gas Lift with a confidence score of 61.3%, correctly matching the documented field selection (Shedid and Yakoot [45]). The leading probability assigned to Gas Lift reflects the model's recognition of the high water cut, moderate GOR, offshore context, and elevated bottomhole temperature as collectively consistent with the Gas Lift operating envelope — in line with the engineering reasoning above.

For Blind test well-2; the Yaran Field well, the model predicted ESP with a confidence score of 58.75%, again matching the documented selection [46]. The moderate confidence level is expected given this well's parameter profile: low GOR and low water cut reduce the discriminative signal that typically separates ESP from Gas Lift in more extreme scenarios, introducing some ambiguity in the probability distribution. Nevertheless, the high production rate, deep well setting, favorable fluid properties, and absence of mechanical constraints provided sufficient evidence for the classifier to assign the highest probability to ESP.

Both predictions are correct, giving a blind well accuracy of 100% (2/2). While the sample size is too small for statistically definitive conclusions, these results demonstrate that the model can correctly apply its learned decision logic to previously unseen wells from geographically and geologically distinct basins, across two different lift methods, and under a wide range of reservoir and fluid conditions. The findings are consistent with the strong cross-validated performance reported in Section 4.1 and support the external validity of the proposed framework.

4.8 Limitations and Future Work

There are some limitations to the results from this study that must be considered, despite being promising. The training dataset was made up of 990 actual well records that have an unbalanced representation of the classes; therefore, it was necessary to use synthetic augmentation by using SMOTE (Synthetic Minority Over-sampling Technique) and domain-informed techniques to create additional records of the minority lift classes in order to have a balanced representation. While these techniques can help to improve the representational

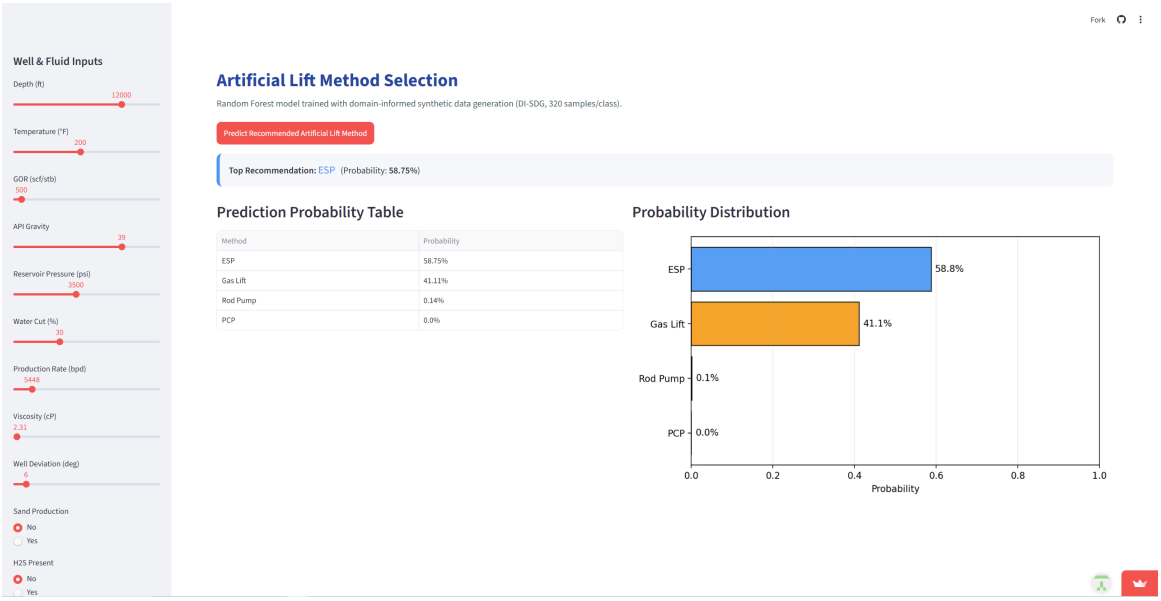

Figure 19. Yaran Field well Streamlit validation.

coverage of the training dataset, they do not provide a substitute for actual field data, nor has the model's generalizability beyond the conditions captured within this research been validated against independent multi-operator datasets.

Currently, the model depends only on static well and reservoir characteristics and does not include the changing nature of the production characteristics. The suitability of a lift method may be changed greatly as a result of a well's production life; therefore, a static snapshot may not take into consideration this temporal evolution of production. In addition, while this framework does deal with the technical feasibility of lifts, it does not take into account economic factors (capital and operating costs) or infrastructure limitations; these latter two considerations are very important factors that will determine the lift selection in practice.

The future work should include increasing the external validation dataset, increasing the number of constraints considered in the techno-economic model for multi-period lift selection, and utilizing XAI techniques to improve the transparency of the model and provide the practitioner with increased confidence in the lift selection model. The incorporation of time-series production data will allow the framework to develop the temporal dimension associated with lift optimization.

5 Conclusion

This study presented a ML based framework for artificial lift method selection that addresses three core challenges in model development: class imbalance, engineering-valid synthetic data generation, and physical interpretability.

A dataset of 990 wells was constructed using 12 domain-informed parameters spanning well geometry, fluid properties, reservoir energy, mechanical constraints, and formation type. The original dataset exhibited a class imbalance ratio of 4.4:1, which was corrected using the proposed Domain-Informed Synthetic Data Generation (DI-SDG) method. Unlike SMOTE, which interpolates purely in feature space and can produce physically implausible parameter combinations, DI-SDG applies feature-specific perturbation strategies, relative perturbation for magnitude-dependent features and absolute perturbation for bounded parameters — while enforcing published screening criteria as hard constraints. The result is a balanced training set of 320 samples per class in which all synthetic wells remain operationally valid.

Random Forest showed the best performance out of all of the models that were evaluated; with a macro-F1 of 0.9255, accuracy of 91.41%, precision of 92.47%, and recall of 92.67% on the test dataset that was set aside. XGBoost and Gradient Boosting were both returned with F1-scores of 0.9125 and 0.9006, respectively, thus affirming that ensemble modeling techniques work best in nonlinear decision space when selecting an appropriate artificial lift method. DI-SDG matched or exceeded the performance of SMOTE across all analyzed classifiers; it is noteworthy to mention the benefit of preserving the characteristics of the physical

and operational information as compared to simply increasing the predictive accuracy on its own.

SHAP value analysis revealed that the five most influential features in the selection process are GOR, production rate, well deviation, fluid viscosity, and well depth; all of which are in agreement with established petroleum engineering screening criteria. Additionally, the presence of nonlinear interactions between GOR–depth as well as GOR–temperature provide evidence that the model has learned multivariate engineering relationships (as opposed to simple thresholding rules) within the training dataset.

Error analysis of the 18 incorrectly classified test samples (9.1% error rate) indicated that most errors were not random, with 15 of the 18 occurring at the ESP–Gas Lift boundary. These cases exhibited a mean GOR of approximately 720 scf/stb (range: 480–1,420 scf/stb) and mean well deviation of 32°, representing operational conditions where both methods are technically viable alternatives and final selection is primarily governed by existing surface infrastructure, capital expenditure constraints, and long-term GOR trajectory — factors not encodable in the 12-feature input space. The model's uncertainty in this zone therefore reflects genuine engineering ambiguity rather than a model deficiency. Of the three remaining errors, two involved PCP wells misclassified as Rod Pump under conditions of intermediate viscosity (mean ˜87 cP), absent sand production, and shallow depth below 3,500 ft, where the absence of a positive sand production flag removes the primary discriminating signal between these two methods; one additional error involved a Rod Pump well misclassified as ESP, representing an isolated boundary case. Wells that were misclassified had a mean prediction confidence of 0.64, significantly lower than 0.82 for correctly classified cases. Therefore, a confidence threshold of 0.75 serves as a practical triage rule, supported by an 18-percentage point separation between these two distributions.

Two blind test wells were then independently verified as a successful extension of the model's ability to generalize correctly to unseen wells with distinct geological environments; both test wells achieved 100% accuracy.

There are still some limitations to the results, particularly because synthetic augmentation cannot fully replace the need for field observations and additional validation with more operators and geological settings is needed. A static reservoir condition does not effectively represent the actual well condition as is true for temporal changes in water cut, GOR, or reservoir pressure occurring over the productive life of a well. In addition, current selection logic does not consider economic factors such as capital and operating costs, infrastructure availability, and issues related to sour-service metallurgy, which constitute a significant gap for practical application at this time.

Future work should focus on including techno-economic optimization, time-based validation of production data with timestamps, field-based validation of production data with timestamps, and the application of explainable AI (XAI) methodologies to continue building practitioner confidence in decision-making. The framework presented here — combining domain-informed data generation, physical mechanism analysis, confidence-based decision triage, and blind well validation — provides a reproducible and interpretable foundation for machine learning applications in production engineering.

Data Availability Statement

The raw production data used in this study are not publicly available due to confidentiality agreements with participating oil companies. However, anonymized summary statistics and aggregated results are provided within the article. The corresponding author may be contacted for reasonable data requests, which will be subject to the aforementioned confidentiality constraints.

Funding

This work was supported without any funding.

Conflicts of Interest

The authors declare no conflicts of interest. For transparency, Author Sarfraz A. Jokhio previously held industry positions with an artificial lift equipment provider (Wood Group ESP) and an operating company (Saudi Aramco); these prior affiliations had no role in the study's design, analysis, or conclusions.

AI Use Statement

The authors declare that Grammarly was used for language editing and grammar correction during the preparation of this manuscript. No generative AI was employed for content generation, data analysis, or interpretation. The authors have reviewed the

manuscript and take full responsibility for the content of this work.

Ethical Approval and Consent to Participate

Not applicable.

References

[1] Clegg, J. D., Bucaram, S. M., & Hein, N. W., Jr. (1993). Recommendations and Comparisons for Selecting Artificial-Lift Methods(includes associated papers 28645 and 29092). *Journal of Petroleum Technology, 45*(12), 1128–1167. [CrossRef]

[2] Lea, J. F., & Nickens, H. V. (1999, March). Selection of artificial lift. In *SPE Oklahoma City Oil and Gas Symposium/Production and Operations Symposium* (pp. SPE-52157). SPE. [CrossRef]

[3] Alemi, M., Jalalifar, H., Kamali, G., & Kalbasi, M. (2010). A prediction to the best artificial lift method selection on the basis of TOPSIS model. *Journal of Petroleum and Gas Engineering, 1*(1), 009–015. https://academicjournals.org/journal/JPGE/article-abstract/DD2D4E82719

[4] Syed, F. I., Alshamsi, M., Dahaghi, A. K., & Neghabhan, S. (2022). Artificial lift system optimization using machine learning applications. *Petroleum, 8*(2), 219–226. [CrossRef]

[5] Cheraghi, Y., Kord, S., & Mashayekhizadeh, V. (2021). Application of machine learning techniques for selecting the most suitable enhanced oil recovery method; challenges and opportunities. *Journal of Petroleum Science and Engineering, 205*, 108761. [CrossRef]

[6] Sathya, R., & Abraham, A. (2013). Comparison of supervised and unsupervised learning algorithms for pattern classification. *International Journal of Advanced Research in Artificial Intelligence, 2*(2), 34-38. [CrossRef]

[7] Mahdi, M. A. A., Amish, M., & Oluyemi, G. (2023). An Artificial Lift Selection Approach Using Machine Learning: A Case Study in Sudan. *Energies, 16*(6), 2853. [CrossRef]

[8] Yakoot, M. S. E., Ragab, A. M. S., & Mahmoud, O. (2021, October). Machine learning application for gas lift performance and well integrity. In *SPE Europec featured at EAGE Conference and Exhibition?* (p. D021S001R008). SPE. [CrossRef]

[9] Khalili, Y., Ahmadi, M., & Moraveji, M. K. (2025). Time-aware predictive maintenance of electrical submersible pumps using catboost ensemble learning and trend-based labeling. *Journal of Petroleum Exploration and Production Technology, 15*(9), 147. [CrossRef]

[10] Ma, F., Altalbawy, F. M., Patel, P., Manjunatha, R., Kalia, R., Formanova, S., ... & Alam, M. M. (2025). Predictive modeling of oil rate for wells under gas lift using machine learning. *Scientific Reports, 15*(1), 27765. [CrossRef]

[11] Ali, J., Ansari, U., Ali, F., Javed, T., & Hullio, I. A. (2026). Application of machine learning for effective screening of enhanced oil recovery methods. *Reservoir Science, 2*(1), 65-80. [CrossRef]

[12] Mohammed, R., Rawashdeh, J., & Abdullah, M. (2020, April). Machine learning with oversampling and undersampling techniques: overview study and experimental results. In *2020 11th international conference on information and communication systems (ICICS)* (pp. 243-248). IEEE. [CrossRef]

[13] Brown, K. E. (1977). *The Technology of Artificial Lift Methods*. PPC Books. https://arks.org/ark:/13960/s2qk5zj2bp9

[14] Takacs, G. (2015). *Sucker-rod pumping handbook: production engineering fundamentals and long-stroke rod pumping*. Gulf Professional Publishing.

[15] Espin, D. A., Gasbarri, S., & Chacin, J. E. (1994, April). Expert system for selection of optimum Artificial Lift method. In *SPE Latin America and Caribbean Petroleum Engineering Conference* (pp. SPE-26967). SPE. [CrossRef]

[16] Heinze, L. R., Winkler, H. W., & Lea, J. F. (1995, April). Decision Tree for selection of Artificial Lift method. In *SPE Oklahoma City Oil and Gas Symposium/Production and Operations Symposium* (pp. SPE-29510). SPE. [CrossRef]

[17] Matthews, C. M., Zahacy, T. A., Alhanati, F. J. S., Skoczylas, P., & Dunn, L. J. (2007). Progressing Cavity Pumping Systems. In L. W. Lake & J. D. Clegg (Eds.), *Production Operations Engineering: IV* (p. 0). Society of Petroleum Engineers (SPE). [CrossRef]

[18] Zheng, A., & Casari, A. (2018). *Feature engineering for machine learning: principles and techniques for data scientists*. " O'Reilly Media, Inc.". https://dl.acm.org/doi/abs/10.5555/3239815

[19] Powers, M. L. (1994). Depth constraint of electric submersible pumps. *SPE Production & Facilities, 9*(02), 137-142. [CrossRef]

[20] Takacs, G. (2009). *Electrical Submersible Pumps Manual: Design, Operations, and Maintenance*. Gulf Professional Publishing.

[21] Noonan, S. G. (2008, October). The Progressing Cavity Pump Operating Envelope: You cannot expand what you don't understand. In *SPE International Thermal Operations and Heavy Oil Symposium* (pp. SPE-117521). SPE. [CrossRef]

[22] Gamboa, J., Aurelio, O., & Sorelys, E. (2003, October). New approach for modeling progressive cavity pumps performance. In *SPE Annual Technical Conference and Exhibition?* (pp. SPE-84137). SPE. [CrossRef]

[23] Alhanati, F. J. S., Solanki, S. C., & Zahacy, T. A. (2001, April). ESP failures: can we talk the same language?. In *SPE Gulf Coast Section Electric Submersible Pumps*

Symposium (pp. SPE-148333). SPE. [CrossRef]

[24] Lea Jr, J. F., & Rowlan, L. (2019). *Gas well deliquification* (3rd ed.). Gulf Professional Publishing. [CrossRef]

[25] Alfaqih, M. R., Ariwibowo, A., & Juliana, C. T. (2016, November). Performance Analysis for Progressive Cavity Pump PCP Production Scenario in Sandy and Heavy Oil Wells. In *SPE Middle East Artificial Lift Conference and Exhibition* (p. D011S003R004). SPE. [CrossRef]

[26] Taheri, A., & Hooshmandkoochi, A. (2006, May). Optimum selection of artificial-lift system for Iranian heavy-oil fields. In *SPE Western Regional Meeting* (pp. SPE-99912). SPE. [CrossRef]

[27] Castro, V., Leite, D., Lemos, D., Marins, J., Pessoa, R., & Magalhães, J. (2015, May). ESP Application on Heavy Oil in Peregrino Field. In *SPE Artificial Lift Conference-Latin America and Caribbean* (pp. SPE-173948). SPE. [CrossRef]

[28] Fakher, S., Khlaifat, A., Hossain, M. E., & Nameer, H. (2021). A comprehensive review of sucker rod pumps' components, diagnostics, mathematical models, and common failures and mitigations. *Journal of Petroleum Exploration and Production Technology, 11*(10), 3815-3839. [CrossRef]

[29] Arnst, B., Morshed, R., & Pond, B. (2021, August 4). *White paper: Reducing rod lift failure in horizontal wells*. Journal of Petroleum Technology. Society of Petroleum Engineers. https://jpt.spe.org/white-paper-reducing-rod-lift-failure-in-horizontal-wells

[30] Waldner, L., Wonitoy, K., Klaczek, W., & Noonan, S. (2012, October). Thermal Performance Testing of a High-Temperature ESP Motor for SAGD Applications. In *SPE Annual Technical Conference and Exhibition?* (pp. SPE-160317). SPE. [CrossRef]

[31] Zhu, H., Zhu, J., Rutter, R., & Zhang, H. Q. (2021). Experimental study on deteriorated performance, vibration, and geometry changes of an electrical submersible pump under sand water flow condition. *Journal of Energy Resources Technology, 143*(8), 082104. [CrossRef]

[32] Zhu, H., Zhu, J., Lin, Z., Zhao, Q., Rutter, R., & Zhang, H. Q. (2021). Performance degradation and wearing of Electrical Submersible Pump (ESP) with gas-liquid-solid flow: Experiments and mechanistic modeling. *Journal of Petroleum Science and Engineering, 200*, 108399. [CrossRef]

[33] NACE International. (2015). *NACE MR0175/ISO 15156 – Petroleum and Natural Gas Industries: Materials for Use in H_2S-Containing Environments in Oil and Gas Production*. NACE International. https://fouladonline.ir/wp-content/uploads/2017/05/NACE-MR-0175-ISO-15156-2015.pdf

[34] ChampionX. (2022). *Sucker Rod Failure Analysis*. ChampionX. https://www.championx.com/contents/NOR_Sucker%20Rod%20Failure%20Analysis_BR_0322.pdf

[35] Al-Khalifa, M., Pessoa Rodrigues, R., & Sinclair, D. (2022). Electrical Submersible Pump Design Enhancements for Hydrogen Sulfide Harsh Environments. *SPE Production & Operations, 37*(04), 603-615. [CrossRef]

[36] Crnogorac, M., Tanasijević, M., Danilović, D., Karović Maričić, V., & Leković, B. (2020). Selection of artificial lift methods: a brief review and new model based on fuzzy logic. *Energies, 13*(7), 1758. [CrossRef]

[37] Reddy, G. T., Reddy, M. P. K., Lakshmanna, K., Kaluri, R., Rajput, D. S., Srivastava, G., & Baker, T. (2020). Analysis of dimensionality reduction techniques on big data. *IEEE Access, 8*, 54776-54788. [CrossRef]

[38] Bourgoyne Jr., A.T., Millheim, K. K., Chenevert, M. E., & Young Jr., F. S. (1986). *Applied Drilling Engineering*. Society of Petroleum Engineers. https://www.scribd.com/document/449344289/Applied-Drilling-Engineering-pdf

[39] Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794). [CrossRef]

[40] Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics, 29*(5), 1189-1232. [CrossRef]

[41] Feurer, M., & Hutter, F. (2019). Hyperparameter optimization. In *Automated machine learning: Methods, systems, challenges* (pp. 3-33). Cham: Springer International Publishing. [CrossRef]

[42] Tanha, J., Abdi, Y., Samadi, N., Razzaghi, N., & Asadpour, M. (2020). Boosting methods for multi-class imbalanced data classification: an experimental review. *Journal of Big data, 7*(1), 70. [CrossRef]

[43] Temizel, C., Canbaz, C. H., Betancourt, D., Ozesen, A., Acar, C., Krishna, S., & Saputelli, L. (2020, October). A comprehensive review and optimization of artificial lift methods in unconventionals. In *SPE Annual Technical Conference and Exhibition?* (p. D041S053R008). SPE. [CrossRef]

[44] Lehman, M. (2004). Progressing cavity pumps in oil and gas production. *World Pumps, 2004*(457), 20-22. [CrossRef]

[45] Shedid, S. A., & Yakoot, M. S. (2016). Simulation study of technical and feasible gas lift performance. *International Journal of Petroleum Science and Technology, 10*(1), 21-44. https://www.researchgate.net/publication/308062978

[46] Janadeleh, M., Ghamarpoor, R., Kadhim Abbood, N., Hosseini, S., Al-Saedi, H., & Hezave, A. (2024). Evaluation and selection of the best artificial lift method for optimal production using PIPESIM software. *Heliyon, 10*(17), e36934. [CrossRef]

[47] Beal, C. (1946). The viscosity of air, water, natural gas, crude oil and its associated gases at oil field temperatures and pressures. *Transactions of the AIME, 165*(1), 94–115. [CrossRef]

Appendix

A DI-SDG Validation Details

A.1 Perturbation Scale Sensitivity Analysis

To validate the choice of perturbation range for DI-SDG, a sensitivity analysis was conducted across seven uniform α values. A fixed Random Forest configuration was used throughout (n_estimators = 200, max_depth = None, max_features = sqrt, min_samples_split = 2, random_state = 42). Results are summarised in Table A1.

The analysis confirms that $\alpha_0 = 0.10$ yields near-optimal macro-F_1 performance among all uniform-α alternatives.

The feature-specific DI-SDG implementation adopted in the main experiments ultimately outperformed every uniform configuration, achieving a macro-F_1 of 0.9427 compared to the best uniform result of 0.9391 at $\alpha_0 = 0.10$. *Note: The macro-F_1 of 0.9255 reported in Table 5 reflects performance under the primary GridSearchCV-optimized hyperparameter configuration used for all six-model comparisons. The macro-F_1 of 0.9427 reported here reflects a separate sensitivity-analysis run using a fixed Random Forest configuration (not hyperparameter-tuned), and is not directly comparable to the main results.*

A.2 SMOTE k-Neighbors Sensitivity Analysis

The sensitivity of SMOTE to the $k_{\text{neighbors}}$ parameter was evaluated across $k \in \{3, 5, 7, 10, 15\}$ using the same fixed Random Forest configuration. Results are presented in Table A2. The analysis confirms that $k = 5$ yields competitive performance and that the SMOTE results reported in the main paper represent the best achievable performance of this method on the current dataset. Both augmentation approaches were subjected to identical fold-stratified GridSearchCV hyperparameter optimization to ensure a fair comparison.

A.3 H_2S Distribution Analysis

Table A3 shows the H_2S prevalence rates by artificial lift (AL) method across the full dataset of 990 wells. A chi-square test of independence confirmed a statistically significant association between H_2S presence and AL method class ($\chi^2 = 32.83, p < 0.001, df = 3$). However, the practical effect size remained small, consistent with the near-uniform prevalence rates observed across all four lift methods. All the four lift methods exhibited H_2S prevalence rates between 15.0% and 41.1%, with substantial overlap across classes.

No class exhibited a distinctly high or consistently low H_2S pattern, which provides statistical support for the near-zero SHAP importance of this variable (mean |SHAP| = 0.0039; ranked 10th out of 12 features).

A.4 Blind Test Well Dataset

Table A4 presents the input feature vectors for the two independent field cases used for external validation of the Random Forest classifier. Both wells were sourced from the peer-reviewed literature and were not involved in any stage of model development. Asterisked values (*) were assumed or derived; full derivations are provided in Appendix B.

B Derivation of Assumed and Inferred Input Values for Blind Test Wells

Blind Test Well 1- Well A [45]

B1 **Well Deviation**
Well deviation is not reported in Shedid & Yakoot (2016) [45]. Therefore, a deviation angle of 65° was assumed, consistent with standard directional offshore completion practices in the Gulf of Suez, where deviated wells are commonly drilled from fixed platform locations to effectively access reservoir targets.

B2 **Sand Production**
Sand production is not reported in Shedid & Yakoot (2016) [45] . An absence of sand production is assumed on the basis of the predominantly consolidated sandstone character of the Nubia Formation and the absence of any reference to sand-related operational challenges in the source publication.

B3 **H_2S Presence**
H_2S presence is not reported in Shedid & Yakoot (2016) [45]. An absence of H_2S is assumed on the basis of the non-sour service character of the Nubia Formation in the Gulf of Suez and the absence of any reference to sour-service metallurgical requirements in the source publication.

Blind Test Well 2 - Yaran Field [46]

B4 **Production Rate**
A single operating production rate is not explicitly stated in Janadeleh et al. (2024) [46]. The production rate was calculated using the standard

Table A1. Sensitivity analysis of uniform perturbation scale (α_0) on Random Forest classification performance.

α_0	Macro-F_1	Accuracy	Description
0.03	0.9114	0.9040	Very conservative; limited synthetic diversity
0.05	0.9321	0.9242	Conservative; narrow perturbation range
0.08	0.9255	0.9141	Moderate; approaching near-optimal zone
0.10	0.9391	0.9242	Baseline — near-optimal; adopted in main experiments
0.15	0.9255	0.9141	Moderate-aggressive; slight performance plateau
0.20	0.9220	0.9091	Aggressive; marginal decline in F_1
0.25	0.9255	0.9141	Most aggressive tested; stable but no gain over baseline

Note: All runs use fixed Random Forest hyperparameters. The feature-specific DI-SDG implementation used in the main experiments achieved a macro-F_1 score of 0.9427, outperforming the best uniform ($\alpha_0 = 0.10$) configuration by 0.0036.

Table A2. SMOTE $k_{\text{neighbors}}$ sensitivity analysis (Random Forest classifier with fixed hyperparameters).

k	Macro-F_1	Accuracy	Description
3	0.9243	0.9141	Fewer neighbors; higher variance in synthetic samples
5	0.9255	0.9141	Default; best F_1 confirmed
7	0.9306	0.9141	More neighbors; smoother interpolation boundary
10	0.9319	0.9141	Higher smoothing; marginal F_1 improvement
15	0.9290	0.9192	Maximum smoothing tested; slight F_1 decline versus $k = 10$

Note: $k = 5$ was used in all main experiments. The marginal differences across k values ($\Delta F_1 < 0.008$) confirm that SMOTE performance is not highly sensitive to this parameter choice for the current dataset.

Table A3. H_2S prevalence by artificial lift method ($n = 990$ wells).

AL Method	H_2S = No (0)	H_2S = Yes (1)	Prevalence (%)
ESP	340 (85.0%)	60 (15.0%)	15.0%
Gas Lift	228 (76.0%)	72 (24.0%)	24.0%
Rod Pump	162 (81.0%)	38 (19.0%)	19.0%
PCP	53 (58.9%)	37 (41.1%)	41.1%
Total	**783 (79.1%)**	**207 (20.9%)**	**20.9%**

Note: Chi-square test of independence: $\chi^2 = 32.83, p < 0.001, df = 3$. Counts were derived from class proportions applied to class totals (ESP = 400, Gas Lift = 300, Rod Pump = 200, PCP = 90). Prevalence rates span 15.0%–41.1%, showing substantial class overlap with no exclusive H_2S signature for any artificial lift method.

Table A4. Independent field cases from the published literature used for external validation of the Random Forest artificial lift classifier.

Parameter	Blind Test Well-1 Well A [45]	Blind Test Well-2 Yaran Field [46]
Well Depth (ft)	12,440	12,000
Reservoir Pressure (psi)	4,900	3,500
Production Rate (STB/day)	3,325	5,600[a]
GOR (scf/STB)	559	500
Water Cut (%)	92	30
Viscosity (cP)	0.55	2.00*
API Gravity (°API)	29.6	39
Well Deviation (degrees)	65*	5*
Bottom-Hole Temp (°F)	270	200
Sand Production	No*	No*
H_2S Presence	No*	No*
Formation Type	Sandstone (Nubia Fm.)	Carbonate (Sarvak/Fahlian Fm.)
Pump Depth (ft)	4,570	N/A
Documented AL Method	Gas Lift	ESP
Model Prediction	Gas Lift – 61.3%	ESP – 58.75%

Assumed or derived values; see Appendix B for full derivations.
[a]Derived from reported Productivity Index; see Appendix B4. N/A = not available in source publication.

Productivity Index relationship:

$$Q = PI \times (P_r - P_{wf}),$$

where $PI = 2$ STB/day/psi (reported), $P_r = 3,500$ psia (reported), and $P_{wf} = 700$ psia (assumed; equal to 20% of P_r, consistent with the aggressive drawdown conditions typical of ESP installations on high-productivity wells):

$$Q = 2\times(3,500-700) = 2\times2,800 = 5,600\,\text{STB/day}.$$

An assumed P_{wf} of 20% of reservoir pressure is consistent with standard ESP design practice, in which ESPs are typically engineered to achieve a flowing bottomhole pressure of 15–25% of reservoir pressure in order to maximize production from high-deliverability reservoirs (Takács, 2009). It is acknowledged that this P_{wf} assumption, grounded in typical ESP operating practice, introduces a degree of circularity in deriving the production rate input, since ESP suitability is partially inferred from an ESP-informed assumption. This limitation is noted for transparency; the well's documented ESP deployment provides independent confirmation that is not contingent on this derivation.

B5 **Viscosity**
Viscosity is not reported in Janadeleh et al. (2024) [46]. A value of 2.00 cP is assumed, consistent with the expected dead oil viscosity of a 39 °API crude at a reservoir temperature of 200 °F, as estimated from the Beal (1946) [47] viscosity–gravity–temperature correlation.

B6 **Well Deviation**
Well deviation is not reported in Janadeleh et al. (2024) [46]. A value of 5° is assumed, consistent with a near-vertical onshore well configuration typical of Yaran Field development practice.

B7 **Sand Production**
Sand production is not reported in Janadeleh et al. (2024) [46]. An absence of sand production is assumed on the basis of the consolidated carbonate character of the Sarvak and Fahlian formations in southwestern Iran and the absence of any reference to sand-related operational challenges in the source publication.

B8 **H_2S Presence**
H_2S presence is not reported in Janadeleh et al. (2024) [46]. An absence of H_2S is assumed on the basis of the non-sour service characterization of the Yaran Field reservoirs and the absence of any reference to sour-service metallurgical requirements in the source publication.

Sohail Nawab is a Ph.D. candidate in Petroleum Engineering (Systems Engineering concentration) at the University of Louisiana at Lafayette, USA, and currently serves as a Graduate Teaching Assistant. He has prior academic experience as a Lecturer at Mehran University of Engineering and Technology (MUET), Pakistan. His research interests include machine learning applications in petroleum engineering, production optimization, and nanoparticles including integrating experimental studies with reservoir simulation and data-driven modeling in Python. (Email: sohailnawabkh@hotmail.com; sohailnawab1@louisiana.edu)

Muhammad Ali is with the Department of Engineering Data Science, University of Houston, USA. His interests include machine learning and data-driven decision support for engineering applications, with emphasis on building robust predictive models and practical feature engineering for real-world applications. (Email: mali68@cougarnet.uh.edu)

Imran A. Hulio received his degree in Petroleum and Natural Gas Engineering from Mehran University of Engineering and Technology (MUET), Jamshoro, Pakistan. His areas of specialization include hydraulic fracturing, reservoir engineering, and production engineering. His primary research interests focus on challenges associated with CO_2 injection and hydraulic fracturing. (Email: imran.hullio@faculty.muet.edu.pk)

Noshad is a Ph.D. candidate in Petroleum Engineering at China University of Petroleum (Beijing), China, with four years of professional experience in drilling operations. His current research focuses on analytically informed machine-learning frameworks for collapse pressure prediction and wellbore stability assessment. His work integrates rock failure criteria, geomechanical constraints, well-log and borehole trajectory data, ensemble learning, uncertainty quantification, and explainable AI to improve safe mud weight window prediction in complex subsurface formations. (Email: alinoshad22@gmail.com)

Ning Liu is an Associate Professor of Petroleum Engineering at the University of Louisiana at Lafayette and is affiliated with the Center for Optimization of Petroleum Systems (COPS). His research areas include petroleum resource management and petroleum technology development, with an emphasis on optimization and data-driven methods for improving petroleum systems performance. (Email: ning.liu@louisiana.edu)

Sarfraz A. Jokhio is an Artificial Lift and Reservoir Management professional with 20+ years of experience in upstream oil and gas industry. He also served as a Professor at Mehran University of Engineering and Technology (MUET) and held reservoir management leadership roles at Saudi Aramco. Earlier, he worked at Schlumberger (SLB) as a Petroleum Reservoir Engineer and at Wood Group ESP as an ESP Application Engineer. (Email: sjokhio2@hotmail.com)