



UNIVERSIDAD ESTATAL PENÍNSULA DE SANTA ELENA

FACULTAD DE CIENCIAS DE LA INGENIERÍA

CARRERA DE INGENIERÍA EN PETRÓLEOS

TEMA:

**“EVALUACIÓN PETROFÍSICA UTILIZANDO MACHINE LEARNING
APLICADO AL CAMPO “X” DEL ORIENTE ECUATORIANO”**

TRABAJO DE INTEGRACIÓN CURRICULAR

AUTOR:

MAYERLY NOEMI BAUTISTA ORTIZ

TUTOR:

ING. FIDEL VLADIMIR CHUCHUCA AGUILAR, MSc

LA LIBERTAD – ECUADOR

2024

UNIVERSIDAD ESTATAL PENÍNSULA DE SANTA ELENA

FACULTAD DE CIENCIAS DE LA INGENIERÍA

CARRERA DE INGENIERÍA EN PETRÓLEOS

TEMA:

**“EVALUACIÓN PETROFÍSICA UTILIZANDO MACHINE LEARNING
APLICADO AL CAMPO “X” DEL ORIENTE ECUATORIANO”**

TRABAJO DE INTEGRACIÓN CURRICULAR

AUTOR:

MAYERLY NOEMI BAUTISTA ORTIZ

TUTOR:

FIDEL VLADIMIR CHUCHUCA AGUILAR, MSc

LA LIBERTAD – ECUADOR

2024

UPSE

TRIBUNAL DE SUSTENTACIÓN



Ing. Marllelis Gutiérrez H, PhD.
DIRECTORA DE CARRERA



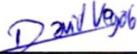
Ing. Fidel Chuchuca, MSc.
DOCENTE TUTOR



Ing. Sadi Iturralde Kure, MSc.
DOCENTE ESPECIALISTA



Ing. Carlos Malavé, MSc
DOCENTE GUÍA DE LA UIC



Ing. David Vega González.
SECRETARIO DEL TRIBUNAL

DEDICATORIA

Dedico este trabajo con gran cariño, principalmente a mi madre y hermana, por haber sido pilares en mi vida, haber estado desde siempre e inspirarme como mujeres, con el ejemplo siempre a más.

A Katherin y Dabor, por su apoyo y amistad incondicional durante gran transcurso de mi vida.

A las familias de Katherin, Jerson, Lisbeth, Angela y aquellos, que me acogieron como un integrante más, en muchos momentos de mi periodo educativo.

Y a mí, por haberlo conseguido, por haber tenido el coraje de soñarlo y alcanzarlo.

CERTIFICADO DE ANTIPLAGIO

En calidad de tutor del trabajo de investigación para titulación del tema “EVALUACIÓN PETROFÍSICA UTILIZANDO MACHINE LEARNING APLICADO AL CAMPO “X” DEL ORIENTE ECUATORIANO” elaborado por la estudiante Mayerly Noemi Bautista Ortiz, egresada de la carrera de Ingeniería en Petróleos, de la Facultad de Ciencias de la Ingeniería, me permito declarar que una vez analizado en el sistema antiplagio COMPILATIO, luego de haber cumplido con los requerimientos exigidos de valoración, la presente tesis, se encuentra con un 6% de la valoración permitida.

 **CERTIFICADO DE ANÁLISIS**
magister

TESIS_EVALUACIÓN PETROFÍSICA ML_MAYERLY BAUTISTA

6%
Textos sospechosos

5% Similitudes
< 1% similitudes entre comillas
2% Idioma no reconocido
0% Textos potencialmente generados por la IA

Nombre del documento: TESIS_EVALUACIÓN PETROFÍSICA ML_MAYERLY BAUTISTA.docx ID del documento: 4b53817e2a42a02de518e3bab4259b00e1454a6d Tamaño del documento original: 133,32 kB	Depositante: FIDEL VLADIMIR CHUCHUCA AGUILAR Fecha de depósito: 10/12/2023 Tipo de carga: interface fecha de fin de análisis: 10/12/2023	Número de palabras: 11.847 Número de caracteres: 71.136
--	---	--

Ubicación de las similitudes en el documento:



Firmado electrónicamente por:
**FIDEL VLADIMIR
CHUCHUCA AGUILAR**

Ing. Fidel Vladimir Chuchuca Aguilar

C.I: 0704789502

DECLARACIÓN DE AUTORÍA

Yo, Mayerly Noemi Bautista Ortiz, declaro bajo juramento que el presente trabajo de titulación denominado "**EVALUACIÓN PETROFÍSICA UTILIZANDO MACHINE LEARNING APLICADO AL CAMPO “X” DEL ORIENTE ECUATORIANO**", no tiene antecedentes de haber sido elaborado en la Facultad de Ciencias de la Ingeniería, Carrera de Ingeniería en Petróleo, lo cual es un trabajo exclusivamente inédito y perteneciente de mi autoría.

Por medio de la presente declaración cedo los derechos de autoría y propiedad intelectual, correspondientes a este trabajo, a la Universidad Estatal Península de Santa Elena, según lo establecido por la ley de propiedad intelectual, por su reglamento y por la normativa institucional vigente

Atentamente,



Mayerly Noemi Bautista Ortiz

Autor de Tesis

C.I. 0850032483

CERTIFICACIÓN DEL TUTOR

Ing. Fidel Vladimir Chuchuca Aguilar, MSc.

TUTOR DE PROYECTO DE INVESTIGACIÓN

Universidad Estatal Península de Santa Elena

En mi calidad de Tutor del presente trabajo “EVALUACIÓN PETROFÍSICA UTILIZANDO MACHINE LEARNING APLICADO AL CAMPO “X” DEL ORIENTE ECUATORIANO” previo a la obtención del Título de Ingeniero en Petróleos elaborado por la señora Mayerly Noemi Bautista Ortiz egresado de la Escuela de Petróleos, Facultad Ciencias de la Ingeniería de la Universidad Estatal Península de Santa Elena, me permito declarar que luego de haber orientado, estudiado y revisado, lo apruebo



Firmado electrónicamente por:
**FIDEL VLADIMIR
CHUCHUCA AGUILAR**

Ing. Fidel Vladimir Chuchuca Aguilar

C.I: 0704789502

CERTIFICADO DE GRAMATOLOGÍA

La Libertad, diciembre 16 del 2023.

Lic. ALEXI JAVIER HERRERA REYES

Magíster En Diseño Y Evaluación

De Modelos Educativos

Certifica:

Que después de revisar el contenido del trabajo de integración curricular en opción al título de **INGENIERO EN PETRÓLEOS** de: **BAUTISTA ORTIZ MAYERLY NOEMI**, cuyo tema es: **“EVALUACIÓN PETROFÍSICA UTILIZANDO MACHINE LEARNING APLICADO AL CAMPO “X” DEL ORIENTE ECUATORIANO**, me permito declarar que el trabajo investigativo se encuentra idóneo y puede ser expuesto ante el jurado respectivo para la defensa del tema en mención.

Es todo cuanto puedo manifestar en honor a la verdad.



Docente de Español A: Literatura
C.I. 0924489255
Reg. Senescyt: 1050-14-86052904
Teléfono: 0962989420
e-mail: alexiherrerareyes@hotmail.com

AGRADECIMIENTO

Agradezco a la Universidad Estatal Península de Santa Elena, por haberme permitido comenzar mi vida profesional en sus instalaciones. A la facultad de ciencias de la ingeniería, que día con día trabajan en mejoras. A la carrera de ingeniería en petróleos, por permitirme ser parte de este grupo de profesionales; y a sus docentes, por haber sido parte de mi formación académica y potenciarla en conocimiento, sobre todo a las mujeres que forman y formaron parte de este grupo, porque como mujer me inspiran a seguir y alcanzar sus pasos en un futuro.

A mi tutor Ing. Fidel Chuchuca, por haber tenido la paciencia de guiarme y corregirme en la elaboración de este trabajo. Al ing. Romel Erazo por haberme abierto las puertas a este tema de investigación.

CONTENIDO

Pág.

TRIBUNAL DE SUSTENTACIÓN	iii
DEDICATORIA	iv
CERTIFICADO DE ANTIPLAGIO	v
DECLARACIÓN DE AUTORÍA	vi
CERTIFICACIÓN DEL TUTOR.....	vii
CERTIFICADO DE GRAMATOLOGÍA	viii
AGRADECIMIENTO.....	ix
CONTENIDO	x
LISTA DE FIGURAS.....	xiii
LISTA DE TABLAS.....	xv
RESUMEN	xvii
ABSTRACT	xviii
CAPÍTULO I: INTRODUCCIÓN	1
1.1 PROBLEMA DE INVESTIGACIÓN	1
1.2 ANTECEDENTES	2
1.3 HIPÓTESIS	4
1.4 OBJETIVOS.....	4
1.4.1 Objetivo general	4
1.4.2 Objetivos específicos.....	5

1.5 ALCANCE	5
1.6 VARIABLES.....	5
1.6.1 Variables independientes.....	5
1.6.2 Variables dependientes.....	6
CAPÍTULO II: MARCO TEÓRICO	7
2.1 MACHINE LEARNING (ML)	7
2.1.1 Aprendizaje supervisado.....	7
2.2 EVALUACIÓN DE FORMACIONES	17
2.3 CONCEPTOS TEÓRICOS SOBRE PETROFÍSICA	17
2.3.1 Litología y tipo de roca.....	17
2.3.2 Volumen de arcilla.....	19
2.3.3 Porosidad	20
2.3.4 Permeabilidad	21
2.3.5 Saturación de fluidos	22
2.3.6 Resistividad	23
2.4 REGISTROS ELÉCTRICOS DE POZO	24
2.4.1 Registros litológicos	25
2.4.2 Registros de resistividad.....	29
2.4.3 Registros de porosidad.....	33
2.5 Medición de reservas del yacimiento	36
2.5.1 Cálculo volumétrico de hidrocarburos	37
CAPÍTULO III: METODOLOGÍA	38
3.1 POBLACIÓN	38
3.2 MUESTRA.....	38
3.3 MODELO METODOLÓGICO PROPUESTO	38
3.3.1 Técnicas e instrumentos de investigación	39

3.4 DESARROLLO DEL MODELO METODOLÓGICO.....	40
3.4.1 Seleccionar los datos	40
3.4.2 Identificar el tipo de problema.....	40
3.4.3 Validación de datos	41
3.4.4 Elección del modelo de regresión.....	44
3.4.5 Determinación de propiedades petrofísicas	45
3.4.6 Cálculo volumétrico de hidrocarburos	49
CAPÍTULO IV: ANÁLISIS Y DISCUSIÓN DE RESULTADOS.....	51
4.1 Evaluación de datos con ML	51
4.1.1 Análisis de la optimización de los tiempos del procesamiento de datos	53
4.2 Evaluación petrofísica de las formaciones	53
4.2.1 Pozo 1	53
4.2.2 Pozo 2	56
4.2.3 Pozo 3	58
4.2.4 Pozo 4	60
4.2.5 Pozo 5	63
4.3 Evaluación de reservas probables.....	65
CAPÍTULO V: CONCLUSIONES Y RECOMENDACIONES.....	66
5.1 Conclusiones.....	66
5.2 Recomendaciones	68
REFERENCIAS BIBLIOGRÁFICAS	69
ANEXO A.....	73
ANEXO B.....	78
ANEXO C.....	82

LISTA DE FIGURAS

Figura 1. Modelamiento con regresión lineal.....	9
Figura 2. Modelamiento con regresión polinomial de grado 2.	9
Figura 3. Modelamiento con regresión polinomial de grado 3.	10
Figura 4. Aplicación de máquina de soporte vectorial.	11
Figura 5. Modelamiento con máquina de soporte vectorial con kernel poly.	12
Figura 6. Modelamiento con máquina de soporte vectorial con kernel rbf.....	12
Figura 7. Ilustración del árbol de decisión.	13
Figura 8. Funcionamiento del modelo árbol de regresión.....	14
Figura 9. Modelamiento con árbol de regresión.....	14
Figura 10. Ajuste del modelo al comportamiento de los datos.	15
Figura 11. Total de datos a dividir en un Split.	16
Figura 12. Relación entre error y complejidad al realizar un split en los datos.	17
Figura 13. distribución de saturación de fluidos en una muestra hipotética de roca yacimiento.	23
Figura 14. Tipos comunes de roca reservorio en una columna estratigráfica.	25
Figura 15. Ejemplo de la desviación SP de la línea base de lutita.	26
Figura 16. Modelo metodológico propuesto.	39
Figura 17. Métodos de selección de algoritmo por la librería scikit-learn. Tomado de (Van den Bossche et al., n.d.).....	40
Figura 18. Datos faltantes en el pozo 1.	41
Figura 19. Descripción del tipo de datos con los que cuenta el pozo 1.....	42
Figura 20. Mapa de calor con correlación entre los registros del pozo 1.....	43
Figura 21. Mapa de calor mostrando correlación de los registros del pozo 1.....	73
Figura 22. Mapa de calor mostrando correlación de los registros del pozo 2.....	74
Figura 23. Mapa de calor mostrando correlación de los registros del pozo 3.....	75
Figura 24. Mapa de calor mostrando correlación de los registros del pozo 4.....	76
Figura 25. Mapa de calor mostrando correlación de los registros del pozo 5.....	77

Figura 26. Modelamiento de datos para registro GR con ML.....	78
Figura 27. Modelamiento de datos para registro SP con ML.....	78
Figura 28. Modelamiento de datos para registro RESD con ML.....	79
Figura 29. Modelamiento de datos para registro RESM con ML.....	79
Figura 30. Modelamiento de datos para registro MNOR y MINV con ML.....	80
Figura 31. Modelamiento de datos para registro PEF con ML	80
Figura 32. Modelamiento de datos para registro NPFI con ML.....	81
Figura 33. Modelamiento de datos para registro RHOB con ML	81

LISTA DE TABLAS

Tabla 1. Lista de las principales rocas sedimentarias.	18
Tabla 2. Calidad de la porosidad.	21
Tabla 3. Permeabilidad en reservorios de petróleo y gas.	21
Tabla 4. Resistividad de algunos fluidos, rocas y minerales.	24
Tabla 5: Factores fotoeléctricos.....	28
Tabla 6. Diferentes coeficientes y exponentes usados para calcular el factor de formación (F).	30
Tabla 7. Profundidad de investigación de los registros de resistividad.....	30
Tabla 8. Velocidades sónicas y tiempo de tránsito del intervalo para diferentes matrices.	34
Tabla 9. Densidades de la matriz de litologías comunes.....	35
Tabla 10. Información por registro para cada uno de los pozos.	41
Tabla 11. Valores para cálculo de reservas posibles	50
Tabla 12. Resultados del modelo de regresión que más se ajusta al comportamiento de los datos	52
Tabla 13. Respuestas de registros litológicos del pozo 1	54
Tabla 14. Resultados de los cálculos de porosidad del pozo 1	55
Tabla 15. Respuestas de registros resistivos del pozo 1	55
Tabla 16. Resultados de saturación de agua del pozo 1	56
Tabla 17. Respuesta de registros litológicos del pozo 2.....	56
Tabla 18. Resultados de cálculos de porosidad del pozo 2.....	57
Tabla 19. Respuesta de registros resistivos del pozo 2.....	57
Tabla 20. Resultados de saturación de agua del pozo 2	58
Tabla 21. Respuestas de registros litológicos del pozo 3	59
Tabla 22. Resultados de cálculos de porosidad del pozo 3.....	59
Tabla 23. Respuesta de registros resistivos del pozo 3.....	60
Tabla 24. Resultados de saturación de agua del pozo 3	60

Tabla 25. Respuesta de registros litológicos del pozo 4.....	61
Tabla 26. Resultados de cálculos de porosidad del pozo 4.....	61
Tabla 27. Respuestas de registros resistivos del pozo 4.....	62
Tabla 28. Resultados de saturación de agua del pozo 4.....	62
Tabla 29. Respuesta de registros litológicos del pozo 5.....	63
Tabla 30. Resultados de cálculos de porosidad del pozo 5.....	64
Tabla 31. Respuesta de registros resistivos del pozo 5.....	64
Tabla 32. Resultados de saturación de agua del pozo 5.....	64
Tabla 33. Cálculos de las reservas probables.....	65
Tabla 34. Resultados de evaluación petrofísica del pozo 1.....	82
Tabla 35. Resultados de evaluación petrofísica del pozo 2.....	83
Tabla 36. Resultados de evaluación petrofísica del pozo 3.....	84
Tabla 37. Resultados de evaluación petrofísica del pozo 4.....	85
Tabla 38. Resultados de evaluación petrofísica del pozo 5.....	86

**“EVALUACIÓN PETROFÍSICA UTILIZANDO MACHINE LEARNING
APLICADO AL CAMPO “X” DEL ORIENTE ECUATORIANO”**

Autor: Mayerly Noemi Bautista Ortiz

Tutor: Fidel Vladimir Chuchuca Aguilar

RESUMEN

El uso de modelos predictivos machine learning puede utilizarse para evaluar el comportamiento de datos y determinar las propiedades petrofísicas. Este trabajo evalúa datos de registros eléctricos de 5 pozos con modelos predictivos, para una evaluación petrofísica y posterior cálculo de reservas probables. La metodología utilizada, consiste en 8 pasos, comenzando por la selección de los datos de los 5 pozos a utilizar de 20 pozos del campo, seguida de identificar el tipo de problema, de regresión, posteriormente validar los datos y seleccionar los modelos a utilizar, se verifica su calidad con el coeficiente de determinación (R^2), procesando e interpretando los resultados de los modelos regresivos para determinar las propiedades petrofísicas y las reservas de las formaciones. Los resultados indican que el modelo que más se ajusta es árbol de regresión, determinando el volumen de arcilla, porosidad y saturación de agua en las formaciones. Este estudio concluye que los modelos regresivos probados en este trabajo no fueron suficientes para determinar el comportamiento de los datos de todos los registros de pozo y las medidas de las propiedades petrofísicas resultaron ser atractivas en las zonas 3, 7 y 10, determinando las reservas probables para estas.

**PALABRAS CLAVE: MACHINE LEARNING / PROPIEDADES
PETROFÍSICAS /**

**PETROPHYSICAL EVALUATION USING MACHINE LEARNING APPLIED
TO THE “X” FIELD IN EASTERN ECUADOR**

Author: Mayerly Noemi Bautista Ortiz

Tutor: Fidel Vladimir Chuchuca Aguilar

ABSTRACT

Using of predictive machine learning models can be used to evaluate data behavior and determine petrophysical properties. This document evaluates electric logs data from 5 wells using predictive models, for a petrophysical evaluation and next calculation of probable reserves. The methodology used, can be described in 8 steps, starting with the selection of the 5 wells to be used out of 20 wells, followed by identifying the type of problem, regression in this case, then validating the data and selecting the models, verifying their quality with coefficient of determination (R^2), processing and interpreting the results of the regression models to determine the petrophysical properties and reserves of the formations. The results indicate, the best fitting model is regression tree, determining the volume of clay, porosity and water saturation in the formations. This study concludes, the regression models tested in this work were not sufficient to determine the data behavior from all wells logs and petrophysical properties measurements were attractive in zones 3, 7 and 10, determining probable reserves for these zones.

KEYWORDS: MACHINE LEARNING / PETROPHYSICAL EVALUATION

CAPÍTULO I: INTRODUCCIÓN

El análisis de las propiedades petrofísicas puede predecir el comportamiento de formaciones. Los registros de pozo sirven para realizar una valoración cuantitativa y cualitativa de las propiedades de la roca y la relación con los fluidos almacenados. Los tipos de registros utilizados son registros litológicos, registros de resistividad y registros de porosidad. Es necesario tiempo y precisión para determinar las propiedades petrofísicas a partir de la lectura de registros eléctricos de pozos.

Este trabajo busca evaluar las formaciones a través de la herramienta machine learning con el propósito de valorar las propiedades petrofísicas optimizando el tiempo, mediante la estructuración de un algoritmo machine learning bajo un aprendizaje supervisado se puede controlar la calidad de estos datos y optimizar el cálculo de las propiedades petrofísicas.

1.1 PROBLEMA DE INVESTIGACIÓN

Durante la corrida de los registros de pozos, pueden presentarse datos anómalos o atípicos por errores inducidos en la medición, problemas con la herramienta, eventos inesperados, interferencias, mala ejecución, cambios en el entorno o geología inesperada.

Luego de la perforación de los pozos, se debe caracterizar las formaciones, identificando los intervalos de profundidad con presencia de hidrocarburos; este trabajo es tedioso y requiere gran cantidad de tiempo, lo que conlleva que, durante el proceso de determinación de las propiedades petrofísicas con un alto error en la medición, exista un procesamiento de datos adicionales, el cual se vuelve complejo su tratamiento.

Los procedimientos tradicionales para el tratamiento de datos de registros de pozo implican muchas horas de trabajo por cada iteración o simulación que se desarrolle para determinar la caracterización petrofísica y el cálculo de reservas.

1.2 ANTECEDENTES

Rogulina et al., (2022) en su trabajo busca determinar y predecir las propiedades de formación de yacimientos para pozos recién perforados basándose en la similitud de las propiedades de los pozos. Comienzan seleccionando los intervalos de pozos cruciales y agregando características para determinar las propiedades petrofísicas. Luego el algoritmo puntúa la similitud entre los pozos a partir de la información obtenida con anterioridad. Aplicando métodos ROC y AUC, usados para verificar la calidad de los modelos en términos de precisión, obtienen una puntuación de 0.824, siendo 1 el modelo perfecto.

Priezzhev et al., (2019) combinan los datos sísmicos y de registros de pozos para estimar las propiedades petrofísicas y petroelásticas usando una comparación analítica de la inversión sísmica con varias técnicas predictivas multivariadas obteniendo que el operador predictivo no lineal, extraído por algoritmo de aprendizaje automático hace posible en ciertos casos estimar las propiedades petrofísicas del yacimiento con mayor precisión y con menos influencia del sesgo de interpretación.

Balz et al., (2015) presentan un método de predicción de las propiedades de reservorio usando clasificación de respuestas sísmicas. El procedimiento comienza generando modelos de rango realista con variaciones estructurales y petrofísicas en el yacimiento, luego se clasifican las trazas sintéticas obtenidas del modelado utilizando redes neuronales de Kohonen. Los resultados fueron un grupo de clases asociados a parámetros estructurales y petrofísicos conocidos, el primer paso del método definió cuales de estos parámetros se pueden extraer de los datos sísmicos y el segundo las facies obtenidas se utilizan para clasificar las trazas sísmicas reales con el fin de encontrar relaciones similares entre las propiedades del yacimiento y la forma de la respuesta sísmica.

Bittar et al., (2021) presentaron un enfoque basado en una red Q profunda para optimizar el reconocimiento de la coincidencia de profundidad de registros de pozo de múltiples pozos. La red aprovecha las capas convolucionales unidimensionales (1D) y extrae características de los registros de entrada. Poniéndolo a prueba, los resultados fueron que después de un extenso entrenamiento el algoritmo aprendió a tomar las mejores decisiones hasta que encuentra una coincidencia.

Lefranc et al. en 2021 desarrollaron un método automatizado para interpretar las estructuras sedimentarias a partir de las imágenes de pozo. Comenzaron creando un conjunto de datos de entrenamiento de imágenes de pozos 2D, luego lo utilizaron para entrenar modelos de redes neuronales profundas, siendo el mejor modelo el RestNet. Probando el modelo obtuvieron que pudo predecir las diferentes estructuras sedimentarias de los registros y resaltar las transiciones con precisión.

En 2018 Gu et al., presentaron una red neuronal con la capacidad de reconocer litofacies complejas, utilizando datos de dos pozos. Obteniendo que la nueva red tiene la capacidad de reconocer litofacies carbonatadas complejas y podría aplicarse a otros esfuerzos geológicos, como análisis de procesos sedimentarios y construcción de un marco de secuencia.

Simoës et al., en 2022 desarrollaron un método para proporcionar a los petrofísicos un conjunto de registros consistentes a través del flujo de trabajo de corrección automática de registros de múltiples pozos, buscando corregir los errores sistemáticos en los registros estándar. Aplicando el método obtuvieron que corrigió el ruido presente en datos de campo y corrigió las mediciones originales en los registros en los que se encontraron características no consistentes.

En 2016 Gi et al., desarrollaron un flujo de trabajo que integra la interpretación de registros acústicos anisotrópicos, el modelado de fracturas en 3D y las redes neuronales para mejorar la precisión y la eficiencia de la predicción de fracturas para lutitas anisotrópicas. Generando una red neuronal con un número limitado de casos de modelado de fracturas en 3D que utiliza propiedades mecánicas de la roca interpretadas a partir de registros sísmicos con modelos acústicos anisotrópicos seleccionados adecuadamente.

Aplicando el modelo y comparándolo con el simulador isotrópico convencional obtuvieron, que ignorar la anisotropía de la lutita lleva a subestimar los anchos de las fracturas y subestimar las contenciones, longitudes y presiones netas de las fracturas, y que la longitud apuntalada efectiva y la productividad de la fractura varían con la posición de los disparos, el volumen inyectado y la tasa de bombeo en las formaciones de lutita.

En 2022 Ahmed et al., utilizaron el machine learning supervisado para predecir registros sísmicos de corte (DTS). Comparando los tres métodos, bosque aleatorio, regresión de árbol de decisión y soporte de regresión vectorial, se mostraron los mejores resultados para el bosque aleatorio. Como resultado obtuvieron que la técnica desarrollada era efectiva en áreas donde faltaran datos y en pozos antiguos que no cuentan con DTS.

En 2022 Adesina & Olusola,, aplicaron el método de regresión forestal aleatoria, usando cuatro pozos para su aplicación. El primer paso fue comprender la distribución de los datos y filtrar valores atípicos, luego usaron una matriz de correlación para comprender la relación entre el corte de agua y la presión de cabeza de la tubería de flujo, tamaño del grano, relación de gasóleo y producción de líquido, posteriormente aplicaron el modelo. Como resultado obtuvieron que el valor de corte de agua tiene una correlación positiva significativa con la fecha de prueba y con la relación gas oil, los valores pronosticados del gráfico de barras y los valores reales mostraron márgenes de error mínimos, lo que indica que se puede confiar en la precisión del modelo.

1.3 HIPÓTESIS

La aplicación de machine learning hace posible determinar las propiedades petrofísicas en 5 pozos del campo “X” del Oriente Ecuatoriano.

1.4 OBJETIVOS

1.4.1 Objetivo general

Evaluar el yacimiento del campo “X” del Oriente Ecuatoriano mediante la aplicación de un algoritmo machine learning, para la optimización de los tiempos de procesamiento de datos.

1.4.2 Objetivos específicos

- Recolectar la información de los registros de pozo
- Organizar la información recolectada en Excel para posteriormente ejecutarlo en el algoritmo
- Distribuir el conjunto de datos, en 20% para datos de entrenamiento y 80% para datos de prueba, a fin de reconocer el comportamiento de los datos
- Determinar el método de medición y clasificación con el que se evaluará la calidad de los datos.
- Verificar la calidad del modelo en términos de precisión con el fin de mejorar el modelo elegido con anterioridad.
- Determinar la profundidad de cada formación con el fin de que el algoritmo distribuya los datos de los registros de pozo ingresados con anterioridad
- Medir las propiedades petrofísicas de las formaciones
- Determinar las reservas del yacimiento.
- Analizar los resultados obtenidos por el algoritmo y compararlos con los de los registros de pozo.

1.5 ALCANCE

El alcance del proyecto se enfoca en determinar las propiedades petrofísicas de 5 pozos del campo “X” del Oriente Ecuatoriano, con información obtenida de registros de pozo; a partir de un algoritmo de machine learning dicha información es analizada a fin de evaluar las propiedades petrofísicas de las formaciones que atraviesan dichos pozos.

1.6 VARIABLES

1.6.1 Variables independientes.

- Datos de registros de pozo
- Algoritmo machine learning

1.6.2 Variables dependientes.

- Porosidad de la formación
- Saturación de los fluidos de la formación
- Resistividad de los fluidos de la formación
- Tipo de fluido de la formación
- Tipo de roca de la formación

CAPÍTULO II: MARCO TEÓRICO

2.1 MACHINE LEARNING (ML)

Machine learning (ML) en petrofísica proporciona un enfoque alternativo basado en datos para realizar algunas tareas; puede asimilar información de grandes bases de datos e inferir relaciones, reglas y conocimientos ocultos.

También, puede utilizarse para ayudar en muchas tareas de interpretación humana que requieren mucho trabajo, como la identificación de datos erróneos, la clasificación de facies y la segmentación de geo-características a partir de datos. (Pandey et al., 2020)

El objetivo de ML es encontrar una respuesta que ayude en la toma de decisiones empresariales.

No siempre es necesario que la respuesta sea perfecta para que funcione, teniendo en cuenta la incertidumbre y representatividad de los datos adquiridos; si la respuesta es buena, podría considerarse en la toma de decisiones. (Xu et al., 2022)

“Los modelos de regresión solamente se aproximan a los datos, generando ecuaciones que permiten predecir qué valor de la variable dependiente se obtendrá incluso para valores de x que nunca antes se han presentado” (Belyadi & Haghghat, 2021).

2.1.1 Aprendizaje supervisado

“Los algoritmos de aprendizaje supervisado automatizan los procesos de toma de decisiones, el usuario proporciona pares de entradas y salidas deseadas, y el algoritmo encuentra la manera de producir la salida deseada dada la entrada”. (Müller & Guido, 2017)

Según Belyadi & Haghghat (2021) los modelos que pueden utilizarse para crear los algoritmos son los siguientes:

- Regresión lineal
- Regresión polinomial
- Regresión logística
- Vecino más próximo a K (KNN)
- Máquina de soporte vectorial (SVM)
- Árbol de decisión
- Bosque aleatorio
- Arboles adicionales
- Potenciación por gradiente
- Refuerzo de gradiente extremo
- Refuerzo de gradiente adaptativo

Los que se usan en este trabajo, se definen a continuación:

A) *Regresión lineal*. Es la más adecuada para trazar una línea a través de un gráfico que muestre una relación lineal entre dos variables; la regresión lineal (*figura 1*) se utiliza cuando existe un alto grado de colinealidad entre las características. Es una de las formas más sencillas y más utilizadas de los algoritmos de regresión ML. La *ecuación 1* es de regresión lineal:

$$y = mx + b \quad (1)$$

Donde:

- m → pendiente de la recta
- b → intercepción
- x → variable independiente
- y → variable dependiente (característica de salida)

“La mayoría de los problemas relacionados con la ingeniería en petróleo tienen relaciones no lineales entre las características de entrada y de salida” (Belyadi & Haghghat, 2021).

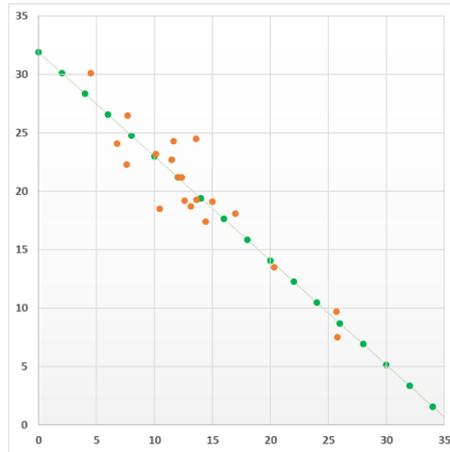


Figura 1. Modelamiento con regresión lineal.

Fuente: Rodríguez & Carecedo, 2022.

B) **Regresión polinomial.** Es adecuada cuando la relación entre características no es lineal; permite disminuir la condición de linealidad manteniendo al mismo tiempo una interpretabilidad alta, añade curvatura al elevar la variable x a diferentes potencias; de esta manera, se pueden conseguir diferentes funciones que representan y se ajustan más a la distribución de los datos. (Rodríguez & Carecedo, 2022)

Se define por la siguiente ecuación:

$$y = a_1x_1 + a_2x_1^2 + b \quad (2)$$

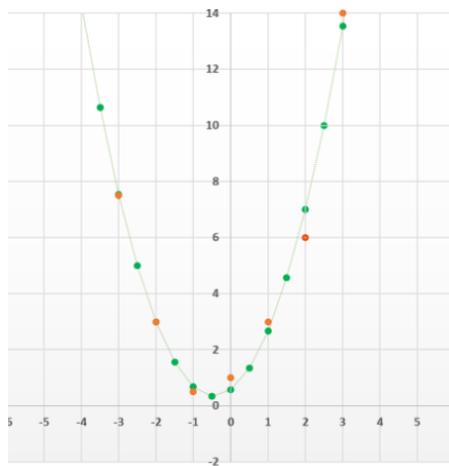


Figura 2. Modelamiento con regresión polinomial de grado 2.

Fuente: Rodríguez & Carecedo, 2022.

En la *figura 2* se ilustra el modelamiento de datos usando regresión polinomial de grado 2.

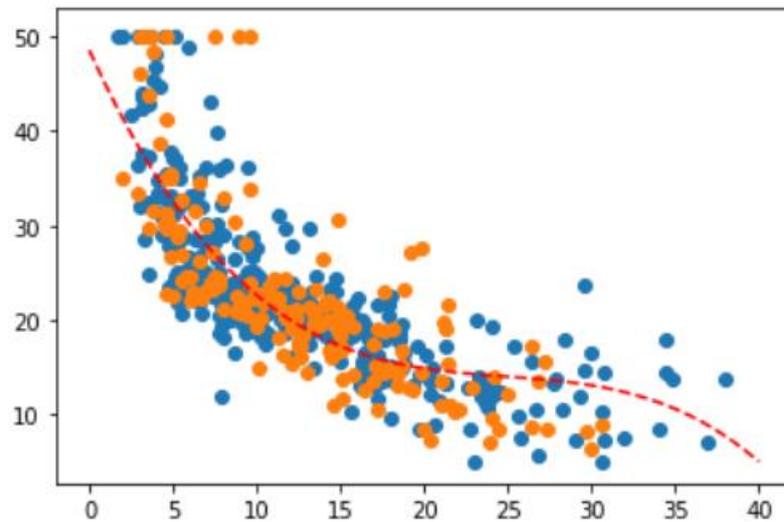


Figura 3. Modelamiento con regresión polinomial de grado 3.

Fuente: Rodríguez & Carecedo, 2022.

En la *figura 3* se muestra el modelamiento de datos usando regresión polinomial de grado 3.

C) *Máquina de soporte vectorial*. Puede utilizarse tanto para problemas de clasificación como de regresión, el objetivo en SVM es maximizar el margen entre vectores de soporte a través de un hiperplano de separación (Cortes & Vapnik, 1995, como se citó en Belyadi & Haghghat, 2021).

Funciona para datos lineales y no-lineales, para explicarlo se puede suponer un conjunto de datos como el que se presenta en la *figura 4*, donde se puede observar que los datos presentan un comportamiento similar a un modelo lineal.

El primer paso para realizarlo es encontrar el modelo lineal que mejor describa el comportamiento de este conjunto de datos, en el contexto de las SVR a la línea se le da el nombre de hiperplano. En problemas donde los datos no son lineales el hiperplano que mejor representa el comportamiento de los datos

será más complejo que una sencilla línea. En este caso la recta será definida por la *ecuación 1*.

Luego, se debe encontrar dos bandas paralelas (ξ en la *figura 4*) al hiperplano donde es necesario que entre ambas cubran la mayor cantidad de datos, estas bandas son conocidas como vectores de soporte.

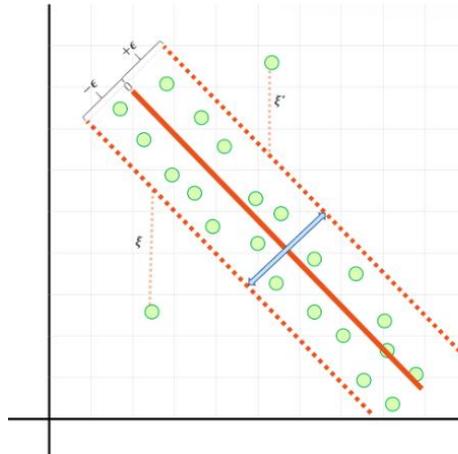


Figura 4. Aplicación de máquina de soporte vectorial.

Fuente: Rodríguez & Carecedo, 2022.

En la *figura 4* se observa cómo el método SVR forma dos fronteras con el fin de definir el comportamiento de los datos (lineal en este caso), también, se visualiza cómo las bandas propuestas no cubren en totalidad a los datos y existen dos datos que están fuera, estos datos son conocidos como errores y son datos que el algoritmo tomará en cuenta para mejorar su desempeño.

En este caso, las bandas paralelas al hiperplano tendrán las siguientes ecuaciones:

$$y = mx + b + \epsilon \quad (3)$$

$$y = mx + b - \epsilon \quad (4)$$

Cuando el comportamiento de los datos no es lineal, se usan “trucos de kernel”. Los trucos de kernel utilizan las características existentes y aplican algunas funciones de transformación para crear nuevas características, existen diferentes funciones Kernel como "linear" (lineal), "poly" (polinomial), "rbf" (función de base radial). (Rodríguez & Carecedo, 2022)

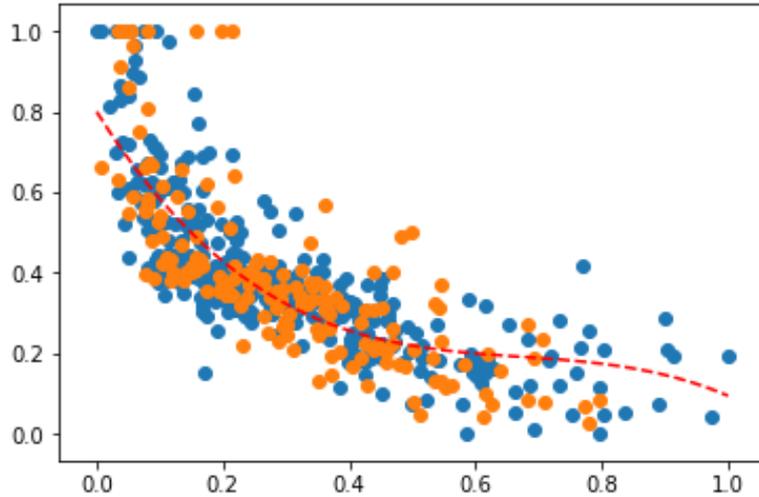


Figura 5. Modelamiento con máquina de soporte vectorial con kernel poly.

Fuente: Rodríguez & Carecedo, 2022.

En la *figura 5* se ilustra el resultado final del modelamiento de datos con máquina de soporte vectorial usando un kernel poly, donde su modelamiento ya no es de una forma lineal como se explicó anteriormente, si no polinomial, es decir los vectores de soporte serán polinomiales y de igual forma el modelamiento.

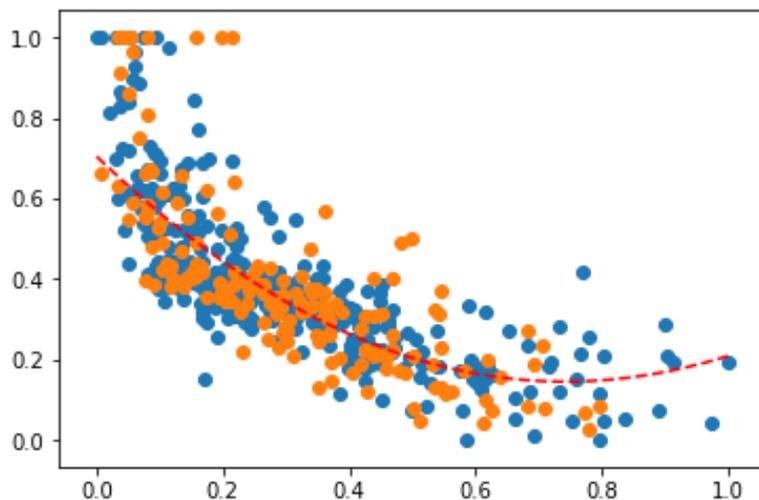


Figura 6. Modelamiento con máquina de soporte vectorial con kernel rbf.

Fuente: Rodríguez & Carecedo, 2022.

En la *figura 6* se indica el modelamiento de datos con máquina de soporte vectorial usando un kernel rbf, un transformador que genera nuevas

características utilizando la distancia entre todos los demás puntos y puntos o centros específicos.

D) *Árbol de regresión*. Se basa en la segmentación en donde el conjunto de datos se divide en diferentes categorías y a los datos que estén dentro de cada intervalo se les asigna directamente un valor establecido en lugar de una etiqueta perteneciente a la categoría donde este ubicado el dato.

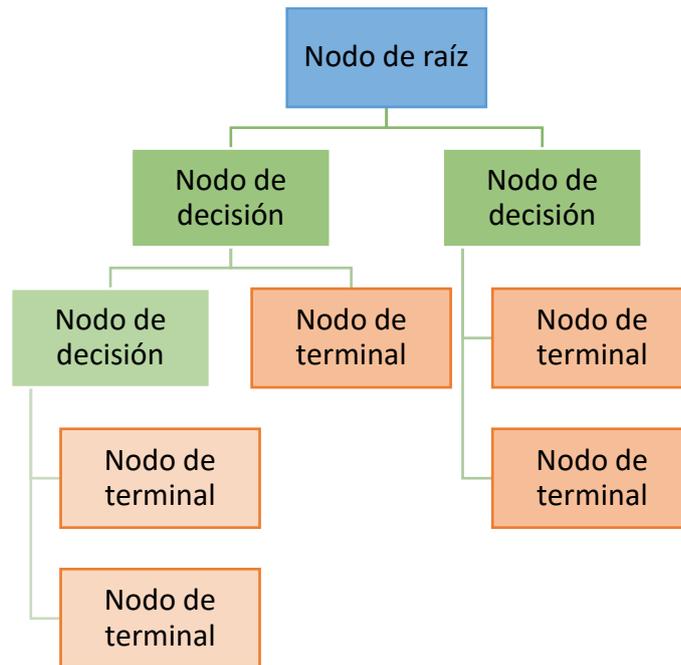


Figura 7. Ilustración del árbol de decisión.

Fuente: Belyadi & Haghghat, 2021.

En la **figura 7** se señala el diagrama de flujo en el que se basa el árbol de regresión de profundidad 2.

El procedimiento para generar un árbol de regresión empieza partiendo la totalidad del intervalo de los datos de la variable independiente x en diferentes intervalos pequeños:

Intervalo 1: $x_1 \leq x \leq x_2$

Intervalo 2: $x_2 \leq x \leq x_3$

Intervalo 3: $x_3 \leq x \leq x_4$

Intervalo 4: $x_4 \leq x \leq x_5$

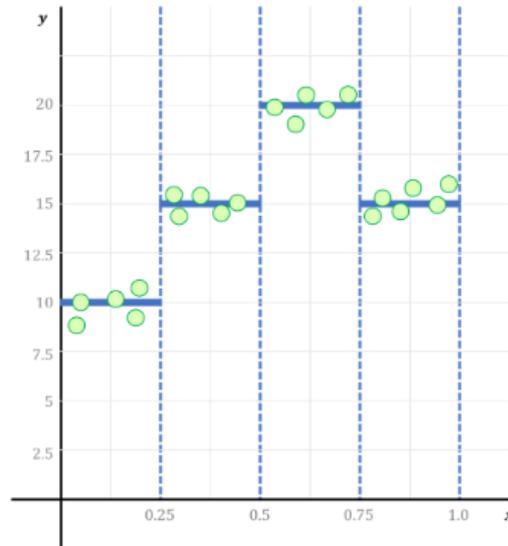


Figura 8. Funcionamiento del modelo árbol de regresión.

Fuente: Rodríguez & Carecedo, 2022.

En la *figura 8* se marcan los intervalos anteriores con la ayuda de las líneas azules punteadas. Y se busca un valor que relacione el conjunto de datos dentro del intervalo (línea azul gruesa) y según esto se puede determinar el valor resultante para ese conjunto de datos.

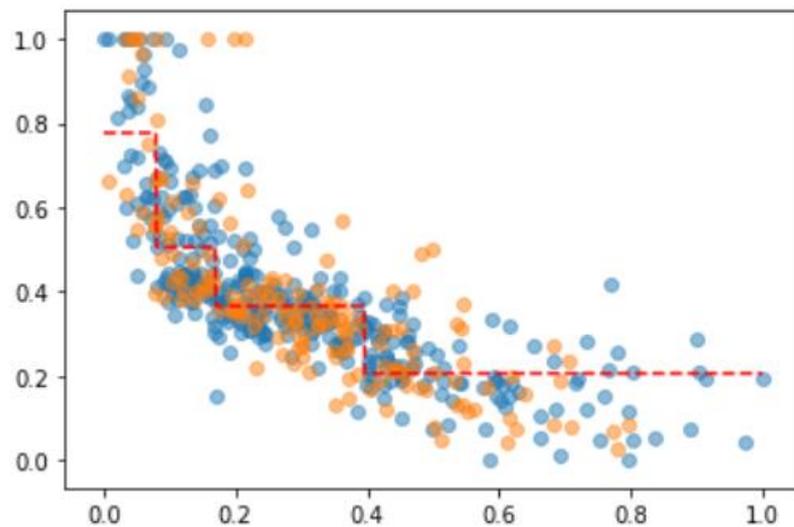


Figura 9. Modelamiento con árbol de regresión.

Fuente: Rodríguez & Carecedo, 2022.

En la *figura 9* se presenta el resultado final del modelamiento de datos con árbol de regresión de profundidad 2.

Acorde a Belyadi & Haghghat, (2021) para evaluar la calidad del modelo de regresión, se usan métricas de error, tales como coeficiente de determinación (R^2), error cuadrático medio (MSE) y error absoluto medio (MAE).

En este trabajo se utiliza el coeficiente de determinación (R^2), es una métrica que indica que tan bien se ajusta el modelo a los datos; su valor varía entre 0 y 1, donde 1 significa que el modelo explica perfectamente la variabilidad de los datos y 0 significa que el modelo no explica ninguna variabilidad. Definido por:

$$\sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (5)$$

Donde:

- y_i → es la salida del punto de muestra i-ésimo.
- \hat{y}_i → es el valor estimado obtenido del modelo.

Teóricamente el valor del modelo evaluado con R^2 es 1 para un ajuste “perfecto” del modelo a los datos y 0 para un modelo “bueno”, un valor negativo indica una pésima predicción.

El valor aceptable para problemas reales es entre 0.7-0.9. Cuando se tiene un modelo con un valor de R^2 cercano a 1 se puede decir que el modelo es “Overfit” (sobre ajuste) y cuando es menor a 0.7 es “Underfit” (bajo ajuste), ninguno de estos dos es el modelo a seleccionar. (Rodríguez & Carecedo, 2022)

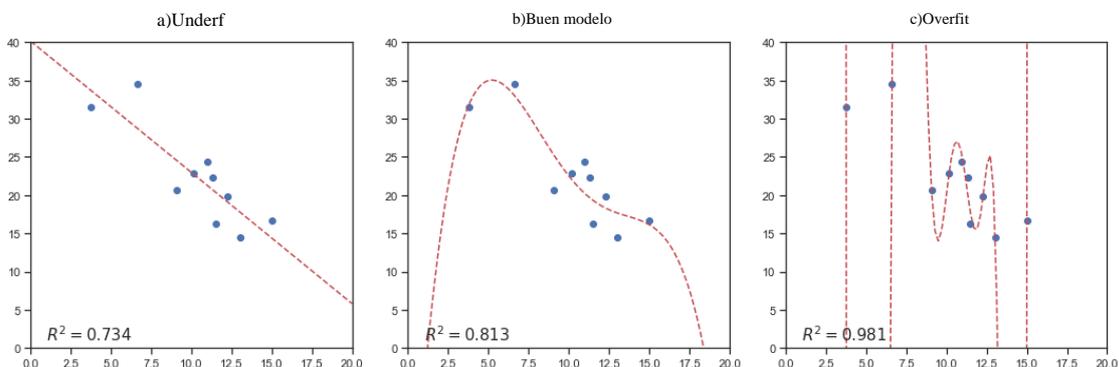


Figura 10. Ajuste del modelo al comportamiento de los datos.

Fuente: Rodríguez & Carecedo, 2022.

En la **figura 10** se indica como se ajusta el modelo al tipo de datos. En la **figura 10a** se usa un modelo lineal para la predicción de datos, y solo 1 dato es interceptado por la línea y los demás son cercanos a esta, teniendo un valor de 0.734 para la métrica de error por R^2 ; en la **figura 10b** se usa un modelo polinomial de grado 4, la curva ya está más cercana al comportamiento de los datos, interceptando a 3, con una métrica de error por R^2 de 0.813; en la **figura 10c** utilizan un modelo polinomial de grado 10, donde la curva del modelo se ajusta a cada uno de los datos, esto es un sobre ajuste, porque el comportamiento del modelo ya no es “general”, si no específico para los datos que tiene y si se quisiera realizar la predicción de un valor que no se encuentra en los datos ($x=5$, por ejemplo) el resultado sería un valor fuera del rango de y (fijarse en la gráfica 10c, el valor de y para un valor $x=5$ según el modelado). Por lo que lo correcto es elegir, un modelo “general” que se ajuste al comportamiento de la mayoría de datos y así mismo logre realizar predicciones dentro de ese rango.

Conforme a Rodríguez & Carecedo (2022) para detectar que se utiliza un modelo adecuado se realiza una división de los datos en conjuntos de entrenamiento y prueba, también llamado split (**figura 11**).

El conjunto de entrenamiento es el subconjunto para entrenar un modelo. El conjunto de prueba es el subconjunto para probar el modelo. El R^2 de ambos conjuntos de datos, deben tener un valor cercano para un buen modelo.

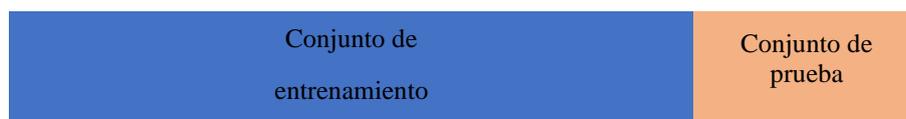


Figura 11. Total de datos a dividir en un Split.

Fuente: Rodríguez & Carecedo, 2022.

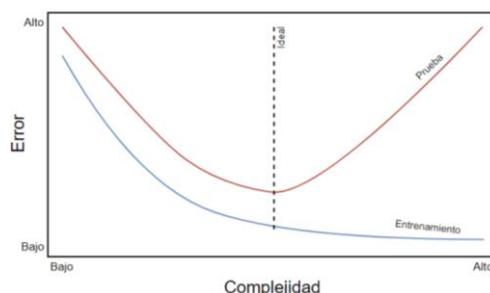


Figura 12. Relación entre error y complejidad al realizar un split en los datos.

Fuente: Rodríguez & Carecedo, 2022.

En la *figura 12* se ilustra que el conjunto de datos de entrenamiento y de prueba deben tener una estrecha relación, al igual que la complejidad y el valor del R^2 .

2.2 EVALUACIÓN DE FORMACIONES

La evaluación de formaciones mediante registros de pozo es el único método que aporta de manera continua las propiedades de la formación a medida que avanza en profundidad, con el objetivo de verificar que el fluido presente es hidrocarburo y cuantificar la cantidad en el medio poroso de la roca reservorio (Camargo, 2005).

2.3 CONCEPTOS TEÓRICOS SOBRE PETROFÍSICA

Según Yang (2017) “la petrofísica es el estudio de las propiedades de las rocas y sus interacciones con los fluidos, tales como, gas, petróleo y agua”. Para ello se definen las propiedades de la roca y de los fluidos a continuación:

2.3.1 Litología y tipo de roca

Según Holstein (2007) el término litología es utilizado para la identificación general de una capa de roca en el subsuelo y se refiere como arenisca, caliza, dolomita, etc.

El tipo de roca es la descripción más detallada de la litología, toma en cuenta las agrupaciones naturales de sistemas de poros que producen propiedades reconocibles utilizadas para predecir propiedades de flujo, volúmenes y saturaciones de fluidos. Es

decir, la litología se centra en los granos, mientras que el tipo de roca se centra en los poros.

Conforme a Tiab & Donaldson (2016) las propiedades físicas de las rocas son consecuencia de su composición mineral.

Las rocas sedimentarias, aproximadamente el 66% de todas las rocas, son las importantes en el estudio de la petrofísica porque la mayor parte del petróleo del mundo se encuentra en rocas sedimentarias porosas. Pueden organizarse según su origen (mecánico, químico y biológico) y composición.

En la *tabla 1* se muestra la lista de las principales rocas sedimentarias.

Tabla 1. Lista de las principales rocas sedimentarias.

Modificado de Tiab & Donaldson, 2016.

Mecanismo	Formación	Composición
	Arenisca	Arenisca Cuarzosa-Granos de cuarzo-Origen deltaico.
Erosión mecánica	Arena friable	Calcáreo-Fragmentos de caliza.
	Clásticas	Clásticos-Granos poco cementados de otras rocas.
	Arena no consolidada	Clásticos-Granos de arena sueltos de otras rocas.
	Silicio	Clásticos-Clásticos compactados, cementados, de grano fino con tamaño de grano inferior a 1/16 bmm.
Erosión química	Conglomerado	Gravas y limos cementados con fango y arena fina
	Lutita	Arcilla-Partículas compactadas de grano fino con tamaño de grano inferior a 1/256. Suele estar laminada en planos de estratificación horizontales definidos.
	Evaporitas	Sales y algunas calizas -Yeso ($\text{CaSO}_4 \cdot 2\text{H}_2\text{O}$) -Anhidrita (CaSO_4) -Caliza (CaCO_3)
Origen biológico	Dolomita ($\text{CaMg}(\text{CO}_3)_2$)	Carbonato-Reacción química con la caliza ($\text{CaMg}(\text{CO}_3)_2$)
	Caliza	Carbonato-Extracción biológica del calcio y precipitación del CaCO_3
	Arrecifes	Carbonato-Restos fósiles de organismos marinos
	Diatomita	Silicatos-Restos silicatados de plantas microscópicas

Según Tiab & Donaldson (2016) los materiales básicos que componen las rocas y su química están asociados a las características petrofísicas de las rocas.

- a) *Arcillas*: Son el producto de la descomposición final de las rocas por meteorización, los depósitos pueden acumularse en la zona donde se formaron originalmente a partir de la meteorización o en los depósitos terminales del movimiento de los vientos, el agua y los glaciares.
- b) *Pizarra*: Se desarrolla en un entorno deposicional a partir de una mezcla de arcilla, limo y material orgánico, todos ellos transportados a la zona por el mar, que son transportados a la zona por el viento, el agua o el hielo (glaciar).
- c) *Arenisca*: Los granos de cuarzo y los fragmentos de rocas mixtas resultantes de la degradación mecánica y química de rocas ígneas, metamórficas y sedimentarias pueden ser transportados a otras zonas y transformados posteriormente en areniscas; después de que los sedimentos sueltos de arena, arcilla, carbonatos, etc., se acumulan en una zona de cuenca sufren entierros por otros sedimentos que se forman encima.
- d) *Carbonatos*: Se forman en ambientes marinos poco profundos, muchos pequeños animales, plantas y bacterias que segregan cal (CaO) viven en aguas poco profundas, sus secreciones y caparazones forman muchas de las rocas carbonatadas.

2.3.2 Volumen de arcilla

Su determinación es un paso esencial en el proceso de evaluación de formaciones, determina la cantidad de arcilla presente en la formación; si no se tiene en cuenta la presencia de arcilla en una formación porosa-permeable, se podría realizar una mala interpretación de la litología y porosidad, por lo tanto, de la saturación de los fluidos presentes en la formación.

Para el cálculo de este parámetro, se pueden utilizar varios indicadores del contenido de arcilla derivado de los registros (Kamel & Mabrouk, 2003).

2.3.3 Porosidad

También llamada porosidad absoluta, “se define como la relación entre el volumen poroso y el volumen total de la roca” (Escobar, 2012).

Expresada como:

$$\begin{aligned} \emptyset &= \frac{V_p}{V_t} \\ &= \frac{\text{Vol de poros interconectados} + \text{Vol de poros no interconectados} + \text{Vol entre poros}}{\text{Vol total}} \quad (6) \end{aligned}$$

Donde:

- \emptyset → Porosidad
- V_p → Volumen poroso
- V_t → Volumen total

Según (Escobar, 2012) “la conexión de los poros, puede clasificarse en”:

- **Porosidad absoluta.** Toma en cuenta el volumen poroso de la roca, sin tomar en cuenta si están interconectados.
- **Porosidad efectiva.** Únicamente toma en cuenta los poros interconectados que permiten el movimiento del fluido.
- **Porosidad no efectiva.** Puede definirse como la diferencias entre la porosidad absoluta y la efectiva, es decir, los poros no interconectados de la roca.

“También, puede verse afectada por el tamaño del grano, la forma del grano, la clasificación, el contenido de arcilla, la compactación y la cementación” (Dandekar, 2013).

En la mayoría de los yacimientos la porosidad varía lateral y verticalmente, encontrándose en un rango de 5 al 20%. (Paris de Ferrer, 2009)

La **tabla 2** muestra la calidad de cada uno de los porcentajes de porosidad. Un yacimiento con una porosidad menor al 5% no es considerada comercial.

Tabla 2. Calidad de la porosidad.

Calidad	Porosidad (%)
Muy buena	>20
Buena	15-20
Moderada	10-15
Pobre	5-10
Muy pobre	<5

2.3.4 Permeabilidad

Según Dandekar (2013) “es la capacidad que tiene la roca de dejar fluir el fluido monofásico a través de ella”

“También es importante mencionar que la permeabilidad de una roca depende de su porosidad efectiva, por lo que también, se ve afectada por el tamaño de grano, empaquetamiento del grano, grado de consolidación y cementación y material cementante” (Tiab & Donaldson, 2016).

La **tabla 3** muestra una clasificación para determinados valores de permeabilidad.

Tabla 3. Permeabilidad en reservorios de petróleo y gas.

Fuente: Tiab & Donaldson, 2016

k (mD)	Clasificación
1 – 15	Mala a regular
15 – 50	Moderada
50 – 250	Buena
250 – 1000	Muy buena
> 1000	Excelente

Según Camargo (2005) se distinguen tres tipos de permeabilidad:

- **Permeabilidad absoluta, K_a .** También conocida únicamente como permeabilidad, es la propiedad de la roca que permite el paso del fluido que la satura al 100%.
- **Permeabilidad efectiva, K_e .** Es la propiedad de la roca para permitir que un fluido fluya a través de ella cuando está saturada por dos o más fluidos no miscibles, por ejemplo, el agua y el petróleo. El agua de formación se adhiere a la

superficie de los poros de la roca, lo que reduce el espacio libre y dificulta el movimiento del petróleo. Esta permeabilidad siempre es menor que la permeabilidad absoluta.

- **Permeabilidad relativa, K_r .** Es la relación entre la permeabilidad efectiva (K_e) y la permeabilidad absoluta (K_a). Se define como:

$$K_r = \frac{K_e}{K_a} \quad (7)$$

2.3.5 Saturación de fluidos

Según Fanchi & Christiansen (2017) el medio poroso de la roca está ocupado por fases fluidas, tales como, gas petróleo y agua; la fracción del volumen de los poros ocupada por determinada fase, se denomina la saturación de esa fase (gas, petróleo o agua).

Conforme a Dandekar (2013) la saturación de agua (S_w) es uno de los factores más importantes a determinar en la evaluación del yacimiento, pues el restante de este valor será el porcentaje de hidrocarburos presente.

Se define a la saturación de fluidos como la relación entre el volumen de una fase fluida (gas, petróleo o agua) en la roca de yacimiento dada y el volumen poroso de la roca:

$$\text{saturación de fluido} = \frac{\text{volumén total de la fase de fluido}}{\text{Volumen poroso total}} \quad (8)$$

Aplicando la ecuación anterior a cada uno de los fluidos presentes, se tiene:

$$S_g = \frac{\text{Volumen de gas}}{\text{Volumen poroso}} \quad (9)$$

$$S_o = \frac{\text{Volumen de petróleo}}{\text{Volumen poroso}} \quad (10)$$

$$S_w = \frac{\text{Volumen de agua}}{\text{Volumen poroso}} \quad (11)$$

Donde:

- S_g → Saturación de gas
- S_o → Saturación de petróleo
- S_w → Saturación de agua

La suma total de los porcentajes debe ser igual a 100% y de las fracciones a 1, esto se muestra en la **figura 13**, siendo:

$$S_g + S_o + S_w = 1 \quad (12)$$

Es importante recordar que las saturaciones de fluidos están estrechamente relacionadas con la permeabilidad relativa.

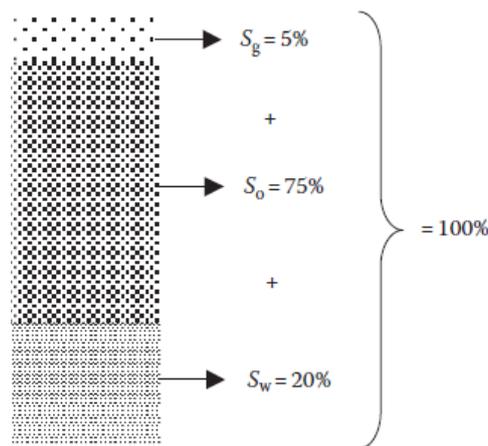


Figura 13. distribución de saturación de fluidos en una muestra hipotética de roca yacimiento.

Fuente: Dandekar, 2013

2.3.6 Resistividad

Según a Asquith & Gibson (1982) “es una medida básica de la saturación de fluidos de un yacimiento y es función de la porosidad, el tipo de fluido y el tipo de roca” y conforme a Camargo (2005) es la resistencia que un material ofrece al paso de la corriente y el parámetro que se pretende medir con las herramientas corridas a pozo abierto, debido a dos razones:

- Las herramientas que miden resistividad tienen profundidad suficiente de investigación, con el fin de medir la resistividad más allá de la zona de invasión.
- Las sondas eléctricas son muy sensibles ante la presencia de hidrocarburos.

La **tabla 4** muestra valores de resistividad de distintos materiales:

Tabla 4. Resistividad de algunos fluidos, rocas y minerales.

Fuente: Camargo, 2005

Materiales	Resistividad ($\Omega \cdot m$)
Fluidos	
Petróleo, gas y aire	Altamente resistivo
Hielo	$10^5 - 10^8$
Agua destilada	$2 * 10^5$
Agua de lluvia	30 – 100
Agua del mar	0.2
Sedimentos no consolidados	
Arcillas	2 – 30
Arenas (varia con la composición)	100 – 5000
Sedimentos consolidados	
Lutita	200 – 500
Areniscas	50 – 1000
Calizas	100 – 50000
Rocas ígneas y metamórficas	10 – 10000

2.4 REGISTROS ELÉCTRICOS DE POZO

Según Asquith & Gibson, (1982) “en la industria petrolera, los registros se refieren a un registro contra profundidad de cualquiera de las características de las formaciones rocosas atravesadas por un aparato de medición en el pozo”

“La imagen que ofrecen los registros es a veces incompleta o distorsionada, pero siempre permanente, continua y objetiva.

Los datos del registro constituyen las características físicas que representan en consecuencia de las condiciones físicas, químicas y biológicas”. (Serra, 1984)

2.4.1 Registros litológicos

Este tipo de registro indica el tipo de roca. La mayoría de acumulaciones de hidrocarburo se da en rocas sedimentarias. En un mismo pozo se pueden encontrar varios tipos de roca (*figura 14*).

Se consideran tres tipos de registros litológicos, el potencial espontáneo (SP), rayos gamma (GR) y fotoeléctrico (PE). (Fanchi & Christiansen, 2017)

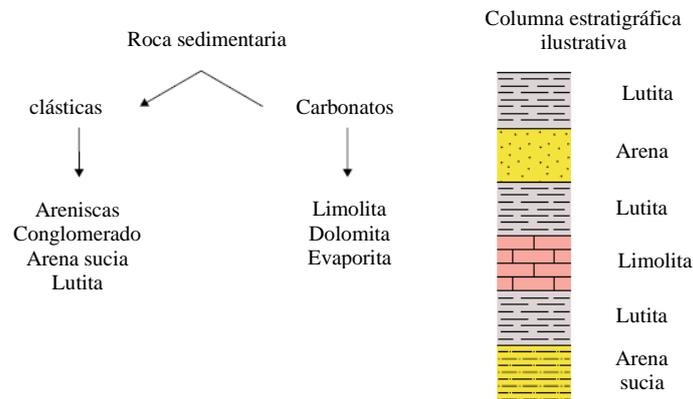


Figura 14. Tipos comunes de roca reservorio en una columna estratigráfica.

Modificado: Fanchi & Christiansen, 2017.

1) *Potencial espontáneo (SP)*. Es usado principalmente para identificar zonas impermeables, como la lutita, y zonas permeables, como la arena. Es de corriente continua (CC), entre el potencial natural de un electrodo móvil en el pozo y el potencial de un electrodo fijo situado en la superficie.

Las respuestas eléctricas son generadas por factores electroquímicos. Los factores electroquímicos del interior del pozo generan la respuesta del registro SP, estos factores electroquímicos se deben a diferencias de salinidad entre el filtrado del lodo (R_{mf}) y la resistividad del agua de formación (R_w) dentro de las formaciones permeables.

Debido a que necesita un fluido conductor en el pozo para que funcione el registro SP, no puede utilizarse en lodos de perforación no conductores (base aceite). Es utilizado para detectar zonas permeables, límites de los estratos permeables, resistividad del agua de formación (R_w) y el volumen de lutita en los estratos permeables.

En este registro está presente el concepto de potencial estático espontáneo (SSP), representa el máximo SP que una formación gruesa, sin lutita, porosa y permeable puede tener para una relación dada entre R_{mf}/R_w ; puede determinarse mediante fórmula o gráfico y es de gran utilidad para determinar valores precisos de R_w y volumen de lutita.

El término Potencial espontáneo pseudoestático (PSP) es el mínimo valor SP que una formación fina, porosa y poco permeable puede tener para la relación R_{mf}/R_w .

El valor de SP que se mide en el pozo está influenciado por el espesor de la resistividad del estrato, la invasión, el diámetro de la perforación y la relación R_{mf}/R_w (**figura 15**). (Asquith & Gibson, 1982)

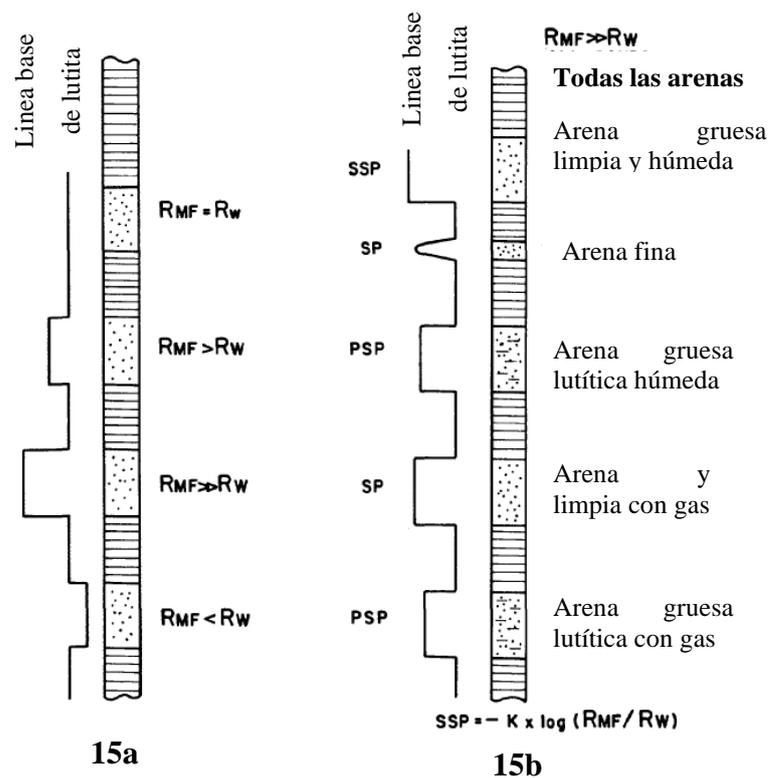


Figura 15. Ejemplo de la desviación SP de la línea base de lutita.

Modificado: Asquith & Gibson, 1982.

En la **figura 15a** se muestra la desviación SP con diferentes resistividades del filtrado de lodo (R_{mf}) y del agua de formación (R_w):

- Cuando $R_{mf} = R_w$ no hay desviación, ni positiva ni negativa de la línea base de lutita.
- Cuando $R_{mf} > R_w$, la línea SP se desvía hacia la izquierda de la línea base de lutita (desviación negativa).
- Cuando R_{mf} es muy superior a R_w , la desviación es proporcionalmente mayor.
- Cuando $R_{mf} < R_w$, la desviación es hacia la derecha de la línea base de lutita, es decir, deflexión positiva.

En la **figura 15b** se muestra desviación SP con resistividad del filtrado del lodo (R_{mf}) mucho mayor que la del agua de formación (R_w). El SSP se presenta en la parte superior del diagrama. Todas las demás deflexiones son menores y de magnitud relativa.

- 2) *Rayos gamma (GR)*. “Estos registros miden la radiactividad natural de las formaciones y, esta medición, puede utilizarse para identificar litologías y correlacionar zonas” (Asquith & Gibson, 1982).

Al igual que el SP, es utilizado para diferenciar zonas porosas permeables potencialmente productoras de hidrocarburos, de las arcillosas no productivas. También, es particularmente útil en pozos con lodo base aceite y revestidos, donde no se puede correr el SP o donde el SP no ofrece buenos resultados porque no existe contraste entre las resistividades de fluido de perforación y el agua de formación, es decir $R_{mf} = R_w$. (Camargo, 2005)

Las areniscas y los carbonatos sin lutita tienen bajas concentraciones de material radiactivo y dan lecturas de rayos gamma bajas. A medida que aumenta el contenido de lutita, aumenta la respuesta del registro de rayos gamma debido a la concentración de material radiactivo en la lutita. Sin embargo, la arenisca limpia (es decir, con bajo contenido de lutita) también puede producir una respuesta de rayos gamma alta si la arenisca contiene feldespatos potásicos, micas, glauconita o aguas ricas en uranio.

Además de su utilidad para identificar litologías y correlación de zonas, los registros de rayos gamma proporcionan información para calcular el volumen de lutita en una arenisca o carbonato. (Asquith & Gibson, 1982)

3) *Fotoeléctrico (PE)*. Según Fanchi & Christiansen (2017) el PE se mide con la absorción de rayos gamma de baja energía por los electrones atómicos orbitales internos, el grado de esta absorción varía con el número atómico del átomo.

La herramienta de registro PE forma parte de las herramientas de registro de densidad, mide la absorción o desvanecimiento de la baja energía gamma, el desvanecimiento de esta energía de rayos gamma se cuantifica como factor fotoeléctrico (PEF); el PEF depende principalmente del número atómico de la formación, los hidrocarburos y el agua contribuyen muy poco al PEF porque sus números atómicos son bajos comparados con los de las rocas de la formación.

Por otro lado, los números atómicos de los diferentes tipos de rocas pueden correlacionarse con el PEF, de modo que el PEF sirve como un buen indicador de la litología o tipo de roca, como se indica en la **Tabla 5**. El PEF no es muy sensible a la porosidad. El registro PE se suele escalar de 0 a 10 en unidades de graneros por electrón (b/e); la profundidad de investigación para la herramienta de registro PE suele ser de 1 a 2 pies.

Tabla 5: Factores fotoeléctricos.

Fuente: Fanchi & Christiansen, 2017

Litología	Aproximación PEF (b/e)
Arenisca (Cuarzo)	2
Dolomita	3
Lutita	3.5
Sal	4.5
Caliza o anhidrita	5

2.4.2 Registros de resistividad

Son utilizados para medir la resistividad de la formación. Los granos de la roca de la formación son no conductores, por lo que la resistividad de la formación está definida por las propiedades eléctricas de los fluidos en el espacio poroso.

Los hidrocarburos son muy resistivos, porque no contienen iones en solución; por el contrario, el agua de formación contiene iones de solución que soportan una corriente eléctrica y su resistividad es relativamente pequeña.

Cumplen con la función de identificar entre formaciones con agua o hidrocarburos. (Holstein, 2007)

Según Asquith & Gibson (1982) conociendo la resistividad del agua de una formación (R_w), su porosidad (ϕ), y un valor para el exponente de cementación (m) (**Tabla 6**) se puede determinar la saturación de agua (S_w) de una formación a partir de la ecuación de Archie:

$$S_w = \left(\frac{F R_w}{R_t} \right)^{1/m} \quad (13)$$

Donde:

- S_w → Saturación de agua
- F → Factor de formación (a/ϕ^m)
- a → Factor de tortuosidad
- m → Exponente de cementación
- R_w → Resistividad del agua de formación
- R_t → resistividad real de la formación medida por un registro de resistividad de lectura profunda
- n → Exponente de saturación (comúnmente 2)

Tabla 6. Diferentes coeficientes y exponentes usados para calcular el factor de formación (F).

Fuente: Asquith & Gibson, 1982.

Factor de formación	Tipo de roca
$F = a/\phi^m$	Relación general Donde: <ul style="list-style-type: none"> • a → Factor de tortuosidad • m → Exponente de cementación • ϕ → Porosidad
$F = 1/\phi^2$	Para carbonatos
$F = 0.81/\phi^2$	Para areniscas consolidadas
$F = 0.62/\phi^{2.15}$	Para arenas no consolidadas
$F = 1.45/\phi^{1.54}$	Para arenas promedio
$F = 1.65/\phi^{1.33}$	Para arenas lutíticas
$F = 1.45/\phi^{1.70}$	Para arenas calcáreas
$F = 0.85/\phi^{2.14}$	Para carbonatos
$F = 2.45/\phi^{1.08}$	Para arenas pliocenas, Sur de California
$F = 1.97/\phi^{1.29}$	Para arenas miocenas, Costa del golfo Texas-Louisiana
$F = 1/\phi^{(2.05-\phi)}$	Para formaciones granulares limpias

Conforme a (Dresser Atlas, 2001, como se citó en Serra, 1984) “Los dos tipos básicos de registros que se utilizan hoy en día y que miden resistividad de la formación son los registros de inducción y de electrodos”.

Según (Asquith & Gibson (1982) “algunos ejemplos de herramientas de resistividad de electrodos son: Normal, Lateral, Laterolog, Microlaterolog, Microlog, Proximity Log, y spherically focused logs”.

En la **tabla 7** se muestra la profundidad de investigación de los registros:

Tabla 7. Profundidad de investigación de los registros de resistividad.

Fuente: Asquith & Gibson, 1982.

Zona lavada (R_{xo})	Zona de invadida (R_i)	Zona virgen (R_t)
Microlog (ML)	Normal corta	Normal larga
Microlaterolog (MLL)	Laterolog-8	Lateral
Registro de proximidad (PL)	Registro esférico enfocado (SFL)	Inducción profunda
Registro micro esférico enfocado (MSFL)	Registro de inducción media	Laterolog profundo

Registro lateralog poco profundo	Lataralog-3
	Lateralog-7
	Inducción

A continuación, se detallan algunos de los registros de resistividad antes mencionados:

- *Registro enfocado de inducción doble:* Según Garduza (2019) “realiza mediciones de resistividad a tres diferentes profundidades de investigación. De esta manera, proporciona información para determinar las resistividades de la zona virgen (Usa el dispositivo R_{ILd}), la zona invadida (Usa el dispositivo R_{ILm}) y la zona de transición”.

Consiste en un dispositivo de inducción de lectura profunda (R_{ILd} que mide R_t), un dispositivo de inducción de lectura media (R_{ILm} que mide R_i) y un Laterolog enfocado de lectura superficial (R_{xo}) que es similar al normal corta.

El registro focalizado de doble inducción se utiliza en formaciones que están profundamente invadidas por filtrado de lodo. (Asquith & Gibson, 1982)

- *Registro lateralog:* Según Holstein (2007) “estos dispositivos están diseñados para minimizar la influencia de los efectos cercanos a la herramienta (es decir, la perforación y la zona invadida) para una medición laterolog profunda, así como para reducir la respuesta de los lechos adyacentes”.

Mide la resistividad real de la formación (R_t) en perforaciones rellenas con lodos salinos (donde $R_{mf} \approx R_w$).

La invasión puede influir en el Laterolog; sin embargo, debido a que la resistividad del filtrado de lodo es aproximadamente igual a la resistividad del agua de formación ($R_{mf} \approx R_w$) cuando se perfora un pozo con lodos a base de agua salada, la invasión no afecta en gran medida a los valores de R_t derivados de un Laterolog. Pero cuando se perfora un pozo con lodos a base de agua dulce (donde $R_{mf} > 3 R_w$), el Laterolog puede verse muy afectado por la invasión, por lo que, no debe utilizarse. (Asquith & Gibson, 1982)

- *Registro Dual Lateralog (DLL)*: “El Dual Lateralog consiste de una lectura profunda (R_i), dispositivo de resistividad (R_{LLd}) y lectura superficial (R_i), dispositivo de resistividad (R_{LLS})” (Asquith & Gibson, 1982).

“En formaciones altamente resistivas, el dual Lateralog es más confiable. Por ejemplo, en los carbonatos de baja porosidad las respuestas de resistividad son muy altas; sin embargo, no es posible realizarlo con lodos no conductivos (base aceite)” (Garduza, 2019).

- *Microlog (ML)*: “Es un dispositivo de resistividad tipo almohadilla que detecta principalmente revoque” (Hilchie, 1978, como se citó en Asquith & Gibson, 1982).

Desde la almohadilla, se realizan dos mediciones de resistividad; una se denomina micro normal y la otra, micro inversa; la micro normal investiga de tres a cuatro pulgadas dentro de la formación (midiendo R_{xo}) y la micro inversa investiga aproximadamente de una a dos pulgadas y mide la resistividad del revoque (R_{mc}).

La detección del revoque Microlog indica que se ha producido una invasión y que la formación es permeable. Las zonas permeables aparecen en el Microlog como separación positiva cuando las curvas micro normales indican una resistividad superior a la de las micro curvas inversas. Las zonas de lutita se indican mediante la ausencia de separación o separación negativa (es decir, micro normal < micro inversa). La separación positiva sólo puede producirse cuando $R_{mc} > R_m > R_{mf}$.

Las zonas no porosas e impermeables tienen altos valores de resistividad en las curvas micro normal y micro inversa. (Asquith & Gibson, 1982)

- *Registro microlateralog (MLL) y de proximidad (PL)*: El Microlateralog (MLL) y el Proximity Log (PL), al igual que el micro-esférico enfocado (MSFL), son registros de electrodos enfocados de tipo almohadilla diseñados para medir la resistividad en la zona lavada (R_{xo}); está diseñado para investigar a mayor profundidad. Por lo que puede utilizarse con lodos de perforación a base de agua dulce donde la capa de lodo es más gruesa.

El Microlaterolog sólo debe utilizarse con lodos de perforación de base salina. (Asquith & Gibson, 1982)

Debido a que el Microlaterolog está fuertemente influenciado por espesores de revoque superiores a de más de ¼ de pulgada sólo debe utilizarse con lodos de perforación a base de agua salada (Hilchie, 1978, como se citó en Asquith & Gibson, 1982).

2.4.3 Registros de porosidad

Según Asquith & Gibson (1982) “estos registros proporcionan información acerca de la porosidad del yacimiento”. Incluye el registro de densidad, registro neutrón y registro sínico:

- *Registro sínico*: Según Holstein (2007) consiste en el tiempo transcurrido para propagar una onda sonora desde una fuente en una herramienta de registro a través de la formación y de vuelta a un receptor en la herramienta, depende de la porosidad de la formación.

“El tiempo de tránsito del intervalo (Δt) en microsegundos por pie, es el recíproco de la velocidad de una onda sonora de compresión en pies por segundo” (Asquith & Gibson, 1982).

Según Wyllie et.al, 1950, como se citó en Asquith & Gibson, 1982 el tiempo de tránsito del intervalo (Δt) depende tanto de la litología como de la porosidad. Por lo tanto, la velocidad de la matriz de una formación (**Tabla 8**) debe ser conocida para derivar la porosidad sínica, ya sea por gráfico o por la siguiente fórmula:

$$\phi_{sonic} = \frac{\Delta t_{log} - \Delta t_{ma}}{\Delta t_f - \Delta t_{ma}} \quad (14)$$

Donde:

- ϕ_{sonic} → Porosidad sínica derivada
- Δt_{ma} → Tiempo de tránsito del intervalo de la matriz

- Δt_{log} → Tiempo de tránsito del intervalo de la formación
- Δt_f → Tiempo de tránsito del intervalo de los fluidos en el pozo (lodo fresco = 189; lodo salado = 185)

La **tabla 8** muestra las constantes usadas en la fórmula de porosidad sónica

Tabla 8. Velocidades sónicas y tiempo de tránsito del intervalo para diferentes matrices.

Fuente: Asquith & Gibson, 1982.

Tipo de roca	V_{ma} (ft/sec)	Δt_{ma} ($\mu sec/ft$)	Δt_{ma} ($\mu sec/ft$) Usados comúnmente
Areniscas	18 000 a 19 500	55.5 a 51	55.5 a 51
Limolitas	21 000 a 23 000	47.6 a 43.5	47.6
Dolomitas	23 000 a 26 000	43.5 a 38.5	43.5
Anhidritas	20 000	50	50
Sal	15 000	66.7	67
Casing (Hierro)	17 500	57	57

- *Registro de densidad:* Es una herramienta, utiliza rayos gamma de energía intermedia emitidos por una fuente radiactiva en la herramienta para inducir la dispersión Compton en la formación.

En este proceso, los fotones de rayos gamma de la herramienta chocan con los electrones de la formación y transfieren parte de su energía a los electrones; el resto de la energía se dispersa en forma de fotones de rayos gamma de menor energía. La herramienta de registro de densidad detecta entonces algunos de estos fotones dispersos.

La tasa de recuento de rayos gamma en la herramienta depende de la densidad de electrones de la formación, que es proporcional a la densidad aparente de la formación ρ_b en gm/cc (Fanchi & Christiansen, 2017).

Según Asquith & Gibson (1982) la densidad aparente de la formación (ρ_b) es una función de la densidad de la matriz, la porosidad y la densidad del fluido en los poros (lodo salino, lodo fresco o hidrocarburos).

Para determinar la densidad de la porosidad, ya sea mediante un gráfico o mediante cálculo, es necesario conocer la densidad de la matriz (**Tabla 9**) y el

tipo de fluido presente en el pozo. La fórmula para calcular la porosidad del registro de densidad es:

$$\phi_{den} = \frac{\rho_{ma} - \rho_b}{\rho_{ma} - \rho_f} \quad (15)$$

Donde:

- ϕ_{den} → Densidad derivada del registro de densidad
- ρ_{ma} → Densidad de la matriz (**Tabla 9**)
- ρ_b → Densidad aparente de la formación
- ρ_f → Densidad del fluido (1.1 lodo salino, 1 lodo fresco y 0.7 gas)

La **tabla 9** muestra las constantes usadas para la fórmula de porosidad del registro de densidad.

Tabla 9. Densidades de la matriz de litologías comunes.

Fuente: Asquith & Gibson, 1982.

Matriz	$\rho_{ma}(gm/cc)$
Arenisca	2.648
Limolitas	2.710
Dolomitas	2.876
Anhidritas	2.977
Sal	2.032

Donde la invasión de una formación es poco profunda, la baja densidad de los hidrocarburos de la formación aumentará la porosidad de densidad (Asquith & Gibson, 1982).

De acuerdo con Garduza (2019) El registro sirve para estimar la densidad del sistema roca-fluido (RHOB) que posteriormente servirá para calcular la porosidad por densidad (DPHI).

Si el registro de densidad es bajo indica alta porosidad, y si es alto indica baja porosidad.

La unidad de medida es g/cm^3 con un rango de valores que va desde 1.95 a 2.95 g/cm^3 , en rocas calizas y de 1.65 a 2.65 gr/cm^3 para terrígenos y siliciclastos.

- *Registro neutrón:* Según Asquith & Gibson (1982) “es un registro de porosidad que mide la concentración de iones de hidrógeno en una formación” y conforme a Garduza (2019) “sirve para estimar la porosidad neutrónica de las rocas (NPHI)”.

En formaciones limpias (libres de lutita) donde la porosidad está llena de agua o petróleo, el registro de neutrones mide la porosidad llena de líquido.

Los neutrones colisionan con los núcleos del material de formación, y dan lugar a que un neutrón pierda parte de su energía. Debido a que el átomo de hidrógeno es casi igual en masa al neutrón, la pérdida máxima de energía ocurre cuando el neutrón colisiona con un átomo de hidrógeno. Por lo tanto, la máxima pérdida de energía está en función de la concentración de hidrógeno de una formación Asquith & Gibson (1982).

2.5 Medición de reservas del yacimiento

Acorde a Escobar (2012) “existen tres métodos para el cálculo de reservas del yacimiento”:

- Calculo volumétrico
- Balance de materia
- Curvas de declinación

En este trabajo, se usará el cálculo volumétrico. Es el método más factible, acorde a la información disponible.

2.5.1 Cálculo volumétrico de hidrocarburos

Según (Escobar, 2012) “se realiza el cálculo volumétrico para el cálculo de petróleo original con la siguiente ecuación”:

$$N = \frac{7758 V_r \phi S_{oi}}{\beta_{oi}} \quad (16)$$

Donde:

- N → Petróleo inicial en sitio, Barr de formación
- V_r → Volumen de la roca, acre-ft
- ϕ → Porosidad, fracción
- S_h → Saturación de hidrocarburo inicial, fracción
- β_{oi} → Factor volumétrico inicial de petróleo, (Barr/Barr de formación)
- 7758 → factor de conversión (Barr/acre-ft)

$$V_r = A h \quad (17)$$

Donde:

- V_r → Volumen de la roca, acre-ft
- A → Área de la roca, acre
- h → Espesor

CAPÍTULO III: METODOLOGÍA

3.1 POBLACIÓN

La población de este trabajo son 20 pozos del campo “X”.

3.2 MUESTRA

La muestra del presente estudio son 5 pozos del campo “X”. Usando un muestreo por conveniencia, no probabilístico se seleccionaron los 5 pozos que presentaban la mayor cantidad de información de registros de pozo, con el fin de que los datos obtenidos sean representativos.

3.3 MODELO METODOLÓGICO PROPUESTO

Teniendo en cuenta las variables dependientes e independientes:

Variable independiente

- Datos de registros de pozo
- Algoritmo machine learning

Variables dependientes

- Porosidad de la formación
- Saturación de los fluidos de la formación
- Resistividad de los fluidos de la formación
- Tipo de fluido de la formación
- Tipo de roca de la formación

- Volumen de arcilla de la formación

El presente trabajo se fundamenta en una investigación descriptiva y correlacional para examinar propiedades petrofísicas de las formaciones de 5 pozos del Campo “X”.

La investigación descriptiva es utilizada para descubrir nuevos hechos y significados, utilizada con el fin de analizar y recopilar información durante la investigación, obteniendo información cualitativa. Con la investigación descriptiva observaremos el funcionamiento del algoritmo, analizando las propiedades petrofísicas de las formaciones y posteriormente examinarlos para comprobar su veracidad. (Grajales, n.d.)

La investigación correlacional pretende encontrar el grado de relación y la manera en que interactúan dos o más variables entre sí.

3.3.1 Técnicas e instrumentos de investigación

Se utilizan los datos de registros de 5 pozos para la creación del algoritmo. Para llevar a cabo la investigación se revisan fuentes teóricas tales como: tesis y papers. Utilizando el siguiente modelo metodológico:

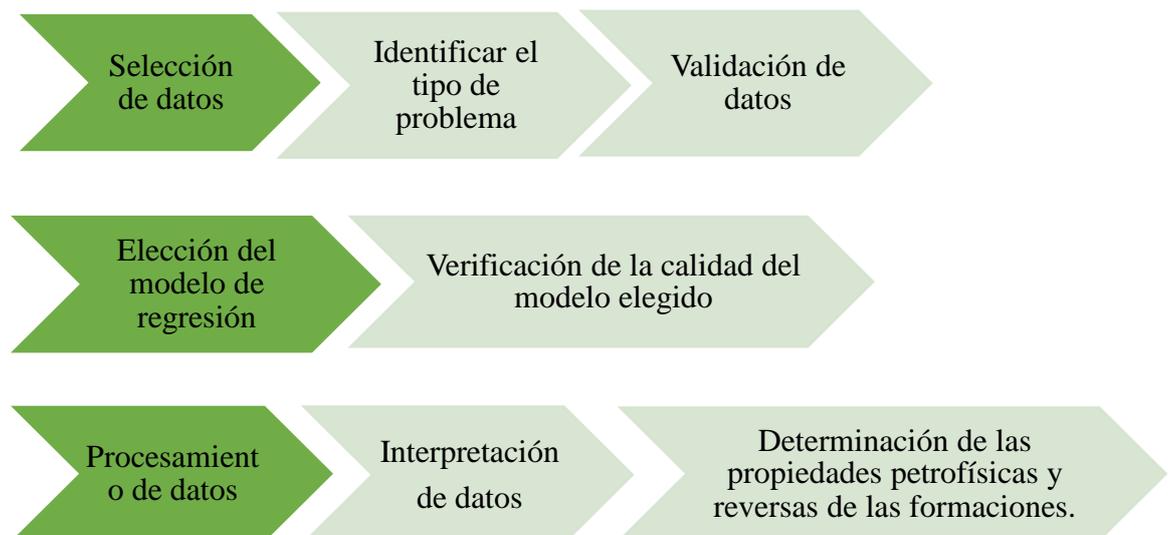


Figura 16. Modelo metodológico propuesto.

3.4 DESARROLLO DEL MODELO METODOLÓGICO

La creación del algoritmo se realiza en Google colab, que mantiene las mismas funciones que Jupyter (si se prefiere, podría usar Jupyter) bajo el lenguaje de programación Python, usando librerías necesarias pertenecientes al lenguaje de programación.

3.4.1 Seleccionar los datos

Se seleccionaron los 5 pozos que presentaban la mayor cantidad de información de registros de pozo.

3.4.2 Identificar el tipo de problema

“Tener una imagen clara del problema que se busca solucionar, entender los datos y saber lo que representan es un paso claro para definir los modelos a utilizar”. (Müller & Guido, 2017)

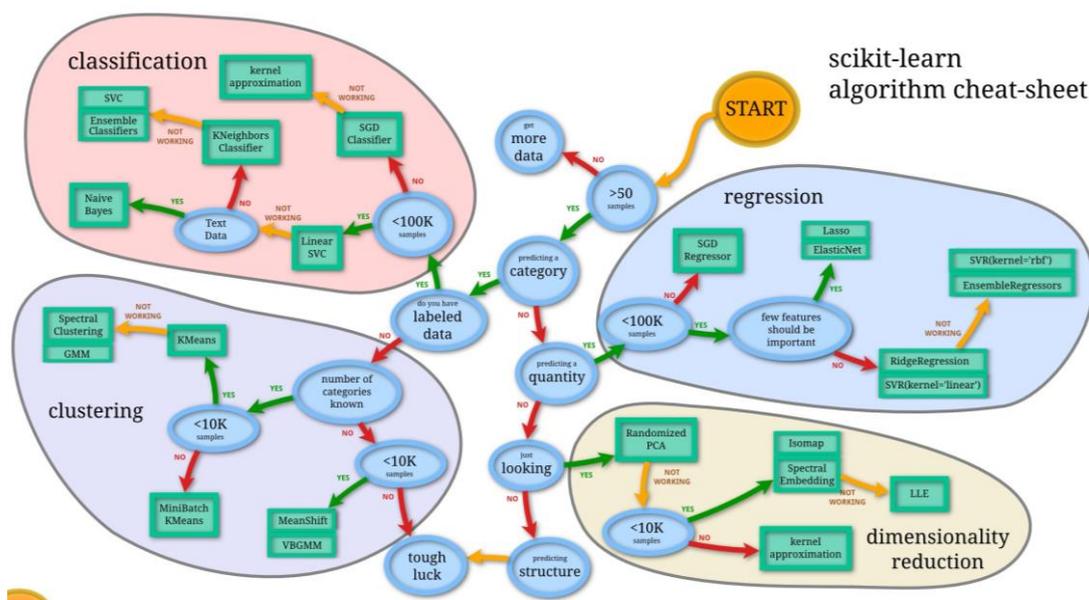


Figura 17. Métodos de selección de algoritmo por la librería scikit-learn. Tomado de (Van den Bossche et al., n.d.)

La *figura 17* muestra un diagrama propuesto por la librería scikit-learn donde ofrece un método para la selección de algoritmos a utilizar dependiendo de las características del problema; al tener más de 50 datos y menos de 100k, numéricos y no categóricos el problema a resolver es de regresión y al tener datos etiquetados será un aprendizaje supervisado.

El código debe ser escrito para cada uno de los pozos, será repetitivo, a menos que se cuente con una característica particular en ese pozo.

3.4.3 Validación de datos

Los pozos cuentan con los siguientes registros, especificados en la *tabla 10*.

Tabla 10. Información por registro para cada uno de los pozos.

Pozo	Registros								
	GR	SP	PEF	RESD	RESM	MINV	MNOR	NPHI	RHOB
1	x	x	x	x	x	x	x	x	x
2	x	x		x	x	x	x		x
3	x	x	x	x	x	x	x	x	x
4	x	x		x	x	x	x		x
5	x	x		x	x	x	x		x

Para cada uno de los pozos se realizó una evaluación del tipo de datos con el que se cuenta, usando la función `.isnull().sum()` para revisar si hay datos faltantes.

```

DEPHT    0
GR       0
SP       0
RESD     0
RESM     0
MINV     0
MNOR     0
PEF      0
NPHI     0
RHOB     0

```

Figura 18. Datos faltantes en el pozo 1.

No hay datos faltantes, pero si, datos con mala lectura (-999.250) por la herramienta de medición. Estos datos se presentan en el final de la última formación (formación H) y se

sustituyen por la media del valor del registro para esa formación. 4 pozos de los 5 pozos presentaron este inconveniente en la última formación. Los valores por mala lectura se sustituyeron por la media del valor de la formación para el registro correspondiente.

Ahora, se puede conocer una breve descripción del tipo de datos y su comportamiento:

	GR	SP	RESD	RESM	MINV	MNOR	PEF	NPHI	RHOB
count	2781.000000	2781.000000	2781.000000	2781.000000	2781.000000	2781.000000	2781.000000	2781.000000	2781.000000
mean	51.147824	-42.764735	17.828205	10.471350	0.051900	1.843288	-50.345913	-69.116344	-51.228565
std	209.351132	146.355252	199.550627	189.333678	50.250294	50.749190	225.815164	254.050474	225.603848
min	-999.250000	-999.250000	-999.250000	-999.250000	-999.250000	-999.250000	-999.250000	-999.250000	-999.250000
25%	64.901001	-23.423000	1.794000	1.836000	1.079000	0.936000	2.736000	0.145000	2.345000
50%	92.482002	-16.770000	3.772000	3.961000	1.869000	2.021000	3.217000	0.230000	2.441990
75%	109.671997	-13.308000	13.758000	15.196000	3.188000	4.094000	3.776000	0.325500	2.513000
max	185.447006	0.570000	1285.913940	1391.800000	11.931000	42.402000	6.300000	0.543900	2.790000

Figura 19. Descripción del tipo de datos con los que cuenta el pozo 1.

La *figura 19* ofrece una rápida descripción de los datos del pozo 1, cantidad de datos, media, moda, mínimo, máximo y porcentajes de cada uno de los registros. Esta descripción de datos se realizó para cada uno de los pozos; en una vista rápida con un mapa de calor de los datos de los registros del pozo 1, se puede notar la correlación entre los datos de los registros.



Figura 20. Mapa de calor con correlación entre los registros del pozo 1.

En la **figura 20** los valores más llamativos son los mayores de 0.5, sin importar su signo y se tiene:

- GR y SP tienen una correlación de 0.51
- GR y NPHI con una correlación de 0.7
- PEF y RHOB cuentan con una correlación de 0.38
- RESD y RESM tienen una correlación 0.92
- RHOB y NPHI con una correlación de -0.48

El mapa de calor (**Revisar Anexo A**) para ojear los mapas de calor de cada uno de los pozos) es una herramienta útil para hacer suposiciones en una primera vista de los datos, estas correlaciones y suposiciones podrían ser aprobadas o refutadas al final. Se realiza

un mapa de calor en cada uno de los pozos, buscando conocer la correlación entre registros.

3.4.4 Elección del modelo de regresión

Para el uso de los modelos de regresión se usará una distribución del conjunto de datos en 20% para datos de entrenamiento y 80% para datos de prueba, para reconocer el comportamiento de los datos.

El siguiente paso es determinar cuál de los métodos de regresión es el adecuado a aplicar para la evaluación del tipo de datos que dispone. Es importante probar los modelos porque si escogiéramos un modelo muy “avanzado”, solo por ser muy avanzado podríamos ocasionar un efecto negativo en el poder de predicción, con un modelo más “general” se obtienen respuestas más reales.

Para cada pozo, una vez realizada la lectura del documento se crea un bucle (for, en este caso) para leer la información de los registros por columna y aplicar los modelos de regresión, En el caso de MINV y MNOR (Lectura del registro microlog para resistividad de la zona somera) se efectúa un condicional (If), buscando realizar el modelo de regresión de forma conjunta con ambas lecturas.

En cada uno de los pozos, en los 4 modelos de regresión se usa $X=Depht$ y $y=el$ registro que se busque evaluar, siendo lo siguiente:

- A) *Regresión lineal.* En el modelado lineal se intenta ajustar una línea a el tipo de datos presente. El modelo hará una predicción lineal que se acople al comportamiento de los datos. Se distribuyen los datos en x y y , según corresponda. Posteriormente, se realiza la distribución de datos del 80/20, 20 para prueba y 80 para entrenamiento. Se grafican los datos para x y y , el modelo de predicción lineal para estos datos. Por último, se realiza el cálculo del R^2 para evaluar el modelo.

- B) *Regresión polinomial.* Para regresión polinomial, se inicia usando una regresión polinomial de grado 2 para comenzar. Si el R^2 de la regresión de grado 2 es menor a 0.7, entonces bajo un condicional se realiza el modelo de regresión polinomial de grado 3 y si este es menor a 0.7, se realiza regresión polinomial de grado 4.

En el modelo de regresión polinomial solo se evaluará hasta grado 4, porque si el modelo de regresión polinomial de grado 4 sigue siendo menor a 0.7, entonces es mejor aplicar otro modelo.

C) *SVR*. Máquina de soporte vectorial se aplica de forma lineal, polinomial y RBF. Para trabajar con el modelo de SVR los datos deben ser ajustados a un mínimo y máximo entre 0 y 1, para se realiza una conversión de datos necesaria usando la función **MinMaxScaler**(), para x y y. Se distribuyen los datos con el Split de 80/20. Después, se realiza regresión con SVR, usando un kernel = rbf; posteriormente, se usa un kernel = poly. Finalmente, se hallan los valores del R^2 para cada uno.

Al final se integran los tres modelos para hacer una comparación en la misma gráfica, mostrando las métricas de error, respectivamente.

D) *Árbol de regresión*. Los árboles de regresión se definen por profundidad, de la profundidad depende el número de salidas. Esta parte del algoritmo empieza usando un árbol de regresión de profundidad 2, si el R^2 es menor a 0.7, se calcula un árbol de regresión de profundidad 3, si el error de este es menor a 0.7, se halla árbol de regresión de profundidad 4 y si el error de este último es menor a 0.7, se encuentra el árbol de regresión de profundidad 5; el modelo de árbol de regresión, también se usa una escala ajustada en valores entre 0 y 1; por último, se realizan las gráficas con el modelo de árbol de regresión con sus respectivas profundidades y métrica de error para cada una.

3.4.5 Determinación de propiedades petrofísicas

Se cuenta con la siguiente información para los pozos 1 y 3 GR, SP, PEF, MINV y MNOR, NPHI, RHOB y para los pozos 2, 4 y 5 no se cuenta con los registros PEF y NPHI. Procediendo a realizar los cálculos de esos pozos sin esa información.

También, se considera los topes de las formaciones para definir cada una de estas.

- **Registros litológicos**. Para el pozo 1 y 3 se analizan los registros SP, GR y PEF. Para los pozos 2, 4 y 5, se analizan el SP y GR.

Para el análisis de los 3 registros litológicos, se realiza lo siguiente:

- (1) Se calcula el valor máximo y mínimo de la lectura del registro en el pozo.
- (2) Se calcula el promedio de la lectura del registro en las formaciones a evaluar.
- (3) Se compara con la mitad del valor máximo de la lectura del registro en el pozo para determinar si es una formación posiblemente rentable.

Y se calcula el volumen de arcilla a partir de GR, considerando las formaciones como areniscas consolidadas:

$$I_{GR} = \frac{GR_{log} - GR_{min}}{GR_{max} - GR_{min}} \quad (18)$$

$$V_{sh} = 0.33 [2^{2I_{GR}} - 1] \quad (19)$$

- **Registros de porosidad.** En el análisis de los registros de porosidad, se halla el promedio del valor registrado por la herramienta para esa formación y se usan las fórmulas para hallar el valor de porosidad según corresponda el registro, es decir:

- Para RHOB, que lee la densidad de la formación se usa la siguiente fórmula para hallar la porosidad:

$$DPHI = \frac{Rhoma - Rhob}{Rhoma - Rhof} \quad (20)$$

Donde:

- *Rhoma* → Constante para cada tipo de roca
- *Rhob* → Densidad del registro
- *Rhof* → Densidad del filtrado de lodo (1, para lodo dulce)

Aquí es de gran utilidad el registro PEF, para determinar la constante de Rhoma.

- En el caso de tener RHOB (lectura por registro por densidad), se calcula DPFI, y si se cuenta con información NPFI (lectura por registro neutrón), se puede estimar TPFI (Porosidad total de la formación):

$$TPFI = \frac{NPFI + DPFI}{2} \quad (21)$$

Donde:

- *DPFI* → Porosidad derivada del registro por densidad
- *TPFI* → Porosidad total de la formación
- *NPFI* → Porosidad derivada del registro neutrón

En los pozos 2 y 4 se usa el valor de porosidad derivado del registro de densidad, al no contar con más registros de porosidad no es posible calcular TPFI.

- Una vez obtenido TPFI, se puede hallar PHIE (Porosidad efectiva) de la formación:

$$PHIE = TPFI * (1 - Vsh) \quad (22)$$

Donde:

- *PHIE* → Porosidad efectiva
- *TPFI* → Porosidad total
- *Vsh* → Volumen de arcilla

- **Registros resistivos.** Utilizando la información que se va obteniendo de las formaciones con el análisis de los registros anteriores, para estudiar los registros resistivos, con el fin de determinar S_w , se realiza lo siguiente:

- Según Asquith & Gibson (1982) basándose en los experimentos de Archie se demostró que la resistividad de una formación llena de agua (R_o), con una resistividad de R_w puede relacionarse mediante un factor de resistividad de la formación (F):

$$R_o = F R_w \quad (23)$$

Donde:

- R_o → Resistividad de la formación 100% saturada de agua
- F → Factor de resistividad de la formación
- R_w → Resistividad del agua de formación

- Los experimentos de Archie también revelaron que los factores de formación pueden relacionarse con la porosidad mediante la siguiente fórmula:

$$F = \frac{a}{\phi^m} \quad (24)$$

Donde:

- F → Factor de resistividad de la formación
- a → Factor de tortuosidad
- m → Exponente de cementación
- ϕ → Porosidad de la formación

En este trabajo se usa la ecuación del factor de formación para arenas lutíticas, por el alto contenido de arcilla de las formaciones:

$$F = \frac{1.65}{\phi^{1.33}} \quad (25)$$

La saturación de agua (S_w) se determina a partir de la resistividad llena de agua (R_o) y la resistividad de la formación (R_t) mediante la siguiente relación:

$$S_w = \sqrt{\frac{R_o}{R_t}} \quad (26)$$

También, representándose por:

$$S_w = \sqrt{\frac{F R_w}{R_t}} \quad (27)$$

Una vez calculada la saturación de agua (S_w), se puede hallar saturación de hidrocarburo para las zonas prospecto:

$$S_h = 1 - S_w \quad (28)$$

3.4.6 Cálculo volumétrico de hidrocarburos

Para calcular las reservas posibles en las formaciones prospecto se usa la siguiente formula:

$$N = \frac{7758 V_r \phi S_h}{\beta_{oi}} \quad (29)$$

Donde:

- N → Petróleo inicial en sitio, Barr de formación
- V_r → Volumen de la roca, acre-ft
- ϕ → Porosidad, fracción
- S_h → Saturación de hidrocarburo inicial, fracción
- β_{oi} → Factor volumétrico inicial de petróleo, (Barr/Barr de formación)
- 7758 → factor de conversión (Barr/acre-ft)

$$V_r = A h$$

Donde:

- V_r → Volumen de la roca, acre-ft
- A → Área de la roca, acre
- h → Espesor

Con los datos especificados a en la siguiente tabla:

Tabla 11. Valores para cálculo de reservas posibles

Formación	Área (Acres)	Espesor (ft)	Boi (BN/BY)
U inferior	36 376	46.5	1.136
T inferior	38 415	35	1.14
H inferior	32 167	115	1.162

CAPÍTULO IV: ANÁLISIS Y DISCUSIÓN DE RESULTADOS

En este capítulo se discuten los resultados obtenidos con el algoritmo del análisis de datos con ML, las propiedades petrofísicas estimadas y las reservas probables.

4.1 Evaluación de datos con ML

En esta sección se discuten los resultados obtenidos del análisis de datos con los modelos regresivos de ML.

El modelo de regresión que más se ajusta al comportamiento de los datos de los 5 pozos es el de árbol de regresión, siendo la única variación, la profundidad de cada uno de los modelos. Se muestra en la **tabla 12** los modelos que cumplieron con el parámetro del R^2 , donde E y P son el valor del R^2 para prueba y entrenamiento, respectivamente. Tree representa el árbol de regresión con su profundidad.

Para los datos del pozo 1 el modelo de árbol de regresión de profundidad 5 es el que mejor se ajusta a su comportamiento, en los registros SP, RESD y RESM; los registros restantes, MINV, MNOR, PEF, NPHI, RHOB tienen un R^2 entre 0.5-0.6 para este mismo modelo, la literatura indica que un valor aceptable de R^2 para el ajuste de un modelo al comportamiento de los datos debe ser entre 0.7-0.9.

En los datos del pozo 2 el modelo de árbol de regresión de profundidad 4 es el que mejor se adapta al comportamiento de los registros SP y RESD; mientras que, para el registro MNOR lo es el árbol de regresión de profundidad 5, para los registros restantes, como lo

son el RESM, MINV y RHOB el valor del R^2 está en un rango de 0.3-0.4, exceptuando al GR, que su valor es 0.6.

Tabla 12. Resultados del modelo de regresión que más se ajusta al comportamiento de los datos

Pozo	Modelo	Registros								
		GR	SP	RESD	RESM	MINV	MNOR	PEF	NPHI	RHOB
1	Tree-2				E: 0,71 P: 0,75					
	Tree-3			E: 0,83 P: 0,90						
	Tree-5	E: 0,61 P: 0,65	E: 0,82 P: 0,80			E: 0,59 P: 0,53	E: 0,60 P: 0,55	E: 0,59 P: 0,59	E: 0,52 P: 0,54	E: 0,53 P: 0,45
2	Tree-4		E: 0,84 P: 0,82	E: 0,77 P: 0,68						
	Tree-5	E: 0,69 P: 0,65			E: 0,32 P: 0,25	E: 0,47 P: 0,45	E: 0,99 P: 0,99			E: 0,32 P: 0,25
3	Tree-5	E: 0,76 P: 0,77	E: 0,78 P: 0,78	E: 0,63 P: 0,57	E: 0,56 P: 0,57	E: 0,43 P: 0,43	E: 0,59 P: 0,52	E: 0,72 P: 0,70	E: 0,60 P: 0,54	E: 0,54 P: 0,47
4	Tree-4		E: 0,74 P: 0,68	E: 0,72 P: 0,55						
	Tree-5	E: 0,66 P: 0,63			E: 0,69 P: 0,74	E: 0,56 P: 0,46	E: 0,61 P: 0,50			E: 0,61 P: 0,50
5	Tree-2	E: 0,89 P: 0,89								
	Tree-3		E: 0,84 P: 0,84							
	Tree-5			E: 0,73 P: 0,69	E: 0,36 P: 0,03	E: 0,53 P: 0,50	E: 0,61 P: 0,58			E: 0,72 P: 0,66

En los datos del pozo 3 el modelo de regresión que mejor se acopla al comportamiento de los registros GR, SP y PEF es el árbol de regresión de profundidad 5. En cambio, para los registros RESD, RESM, MNOR, NPHI y RHOB los valores de R^2 se encuentran entre 0.5-0.6.

En los datos del pozo 4 el modelo de árbol de regresión es el que mejor se ajusta al comportamiento de los registros SP y RESM; por el contrario, para los restantes GR, MINV, MNOR y RHOB cuentan con un valor de R^2 entre 0.5-0.6.

En el pozo 5 el modelo que se ajusta al comportamiento de los datos del registro GR es el de árbol de regresión de profundidad 2, para el SP lo es el de árbol de regresión de profundidad 3 y para RESD y RHOB lo es el árbol de regresión de profundidad 5; en tanto a los registros restantes, MINV y MNOR el R^2 toma un valor de 0.5-0.6 y para RESM un valor de 0.3.

Los modelos lineales, polinomiales hasta grado 4 y máquinas de soporte vectorial no se ajustan al comportamiento de los datos de los 5 pozos; al ser el modelo de árbol de regresión, variando por su profundidad, el que más se ajusta al modelamiento de los datos, nos indica la tendencia de estos a dividirse por segmentos, partiendo a los mismos en pequeños intervalos.

En el *Anexo B*, del capítulo de anexos de este trabajo se muestra algunos de los modelamientos del pozo 1.

4.1.1 Análisis de la optimización de los tiempos del procesamiento de datos

En la optimización de los tiempos de procesamiento de datos, tomo mes y medio la elaboración del código. Al no lograr encontrar un modelo de machine learning que se ajuste al comportamiento de ciertos datos de los pozos, no fue posible la optimización del tiempo del procesamiento de los mismos.

4.2 Evaluación petrofísica de las formaciones

En este apartado se presentan y discuten los parámetros petrofísicos obtenidos de la evaluación petrofísica para los 5 pozos.

4.2.1 Pozo 1

Los resultados obtenidos para el pozo 1, considerando cada uno de los apartados, son los siguientes:

- a) *Litología*. La mayoría de las zonas de este pozo muestran buenas respuestas de GR, 7 zonas (1, 2, 3, 6, 7, 9 y 10) presentan un valor por debajo de la media (92.72), 2 zonas (0, 5) presentan valores ligeramente por encima de la media; y, finalmente 2 zonas (4 y 8) presentan respuestas mayores a 100.

La respuesta del registro SP, son declinaciones negativas en todas las zonas, 3 zonas (3, 7 y 10) presentan las principales declinaciones del pozo, lo que indica posiblemente una zona con alta presencia de permeabilidad.

El registro PEF es de gran utilidad para etiquetar la litología de las formaciones, determinando 4 zonas (0, 2, 6 y 9) como dolomitas, 2 zonas (1 y 5) como calizas, y 3 zonas (3, 7 y 10) como arenas, por ultimo las 2 zonas (4 y 8) como lutitas.

Tabla 13. Respuestas de registros litológicos del pozo 1

ZONAS	Formación	Intervalos	Espesor(ft)	GR	SP	PEF
Zona 0	...	8750-9385	635	95,941	-14,765	3,257
Zona 1	CALIZA-A	9385-9489	104	72,383	-16,932	4,523
Zona 2	U SUPERIOR	9489-9571	82	80,102	-20,54	3,467
Zona 3	U INFERIOR	9571-9617.5	46,5	70,235	-52,806	2,184
Zona 4	...	9617.5-9715	97,5	101,39	-14,645	3,699
Zona 5	CALIZA B	9715-9756	41	92,849	-20,095	3,935
Zona 6	T SUPERIOR	9756-9810	54	77,336	-31,732	3,561
Zona 7	T INFERIOR	9810-9845	35	59,233	-70,889	2,267
Zona 8	...	9845.5-9986	141	135,44	-16,668	3,535
Zona 9	H SUPERIOR	9986-10025	39	80,214	-18,195	3,152
Zona 10	H INFERIOR	10025-10140	115	36,138	-38,271	2,371

- b) **Porosidad.** El volumen de arcilla, de 2 zonas (4 y 8) se encuentra por encima de 0.3, de 5 zonas (0, 2, 5, 6 y 9) se encuentra en un rango de 0.2-0.3, de 2 zonas (1 y 2) con respuestas de 0.18-0.19, y finalmente la zona 10 cuenta con un valor relativamente bajo de 0.043.

La porosidad total de 5 zonas (0, 2, 4, 6 y 8) está por encima de 0.20, considerándose “muy buena” según Paris de Ferrer, 2009, de las 5 zonas (1, 3, 5, 7 y 9) se encuentran en un rango de 0.15-0.20 “buena” y la zona restante cuenta con una porosidad de 0.13 considerada “moderada”.

La porosidad efectiva de 5 zonas (0, 2, 4, 6 y p) se encuentra en un rango de 0.15-0.20 y la de las 6 zonas restantes (1, 3, 5, 7, 8 y 10) se encuentra entre 0.10-0.15.

Tabla 14. Resultados de los cálculos de porosidad del pozo 1

ZONAS	Formación	Intervalos	Espesor(ft)	VSH	PHIE	TPHI
Zona 0	...	8750-9385	635	0,297	0,181	0,257
Zona 1	CALIZA-A	9385-9489	104	0,185	0,138	0,17
Zona 2	U SUPERIOR	9489-9571	82	0,219	0,195	0,249
Zona 3	U INFERIOR	9571-9617.5	46,5	0,194	0,133	0,165
Zona 4	...	9617.5-9715	97,5	0,327	0,201	0,298
Zona 5	CALIZA B	9715-9756	41	0,28	0,138	0,191
Zona 6	T SUPERIOR	9756-9810	54	0,204	0,179	0,225
Zona 7	T INFERIOR	9810-9845	35	0,143	0,136	0,159
Zona 8	...	9845.5-9986	141	0,562	0,105	0,239
Zona 9	H SUPERIOR	9986-10025	39	0,229	0,157	0,203
Zona 10	H INFERIOR	10025-10140	115	0,043	0,125	0,13

c) **Resistividad.** Los valores de MNOR y MINV, donde $MNOR > MINV$, nos indica permeabilidad en la zona. Esta premisa se cumple en todas las zonas.

Los valores más llamativos del registro RESM y RESD son en la zona 3 y 10, con respuestas de resistividad mayores a 100, señalando que son zonas posiblemente con presencia de hidrocarburo.

Tabla 15. Respuestas de registros resistivos del pozo 1

ZONAS	Formación	Intervalos	Espesor(ft)	MINV	MNOR	RESM	RESD
Zona 0	...	8750-9385	635	2,039	3,435	9,745	11,657
Zona 1	CALIZA-A	9385-9489	104	5,241	10,27	20,59	23,94
Zona 2	U SUPERIOR	9489-9571	82	2,519	4,197	11,56	14,434
Zona 3	U INFERIOR	9571-9617.5	46,5	2,38	3,914	101,4	116,35
Zona 4	...	9617.5-9715	97,5	2,132	2,674	5,512	4,3984
Zona 5	CALIZA B	9715-9756	41	4,178	6,105	17,07	20,815
Zona 6	T SUPERIOR	9756-9810	54	2,26	3,074	8,661	7,3474
Zona 7	T INFERIOR	9810-9845	35	2,465	5,718	21,81	22,593
Zona 8	...	9845.5-9986	141	2,782	3,581	10,69	8,442
Zona 9	H SUPERIOR	9986-10025	39	3,323	5,635	33,91	37,738
Zona 10	H INFERIOR	10025-10140	115	2,705	6,105	308,1	391,69

d) **Presencia de hidrocarburo.** Los valores de saturación de agua más llamativos, son los de la zona 3, 9 y 10, con valores menores a 0.30, donde presumiblemente el restante sea hidrocarburo.

Tabla 16. Resultados de saturación de agua del pozo 1

ZONAS	Formación	Intervalos	Espesor(ft)	Sw	Sh
Zona 0	...	8750-9385	635	0,614	...
Zona 1	CALIZA-A	9385-9489	104	0,429	...
Zona 2	U SUPERIOR	9489-9571	82	0,552	...
Zona 3	U INFERIOR	9571-9617.5	46,5	0,194	0,806
Zona 4	...	9617.5-9715	97,5	1	...
Zona 5	CALIZA B	9715-9756	41	0,46	...
Zona 6	T SUPERIOR	9756-9810	54	0,974	...
Zona 7	T INFERIOR	9810-9845	35	0,441	...
Zona 8	...	9845.5-9986	141	0,722	...
Zona 9	H SUPERIOR	9986-10025	39	0,298	0,702
Zona 10	H INFERIOR	10025-10140	115	0,09	0,91

Considerando todo lo relatado, las zonas prospecto de este pozo bajo el análisis de este trabajo son las zonas 3 y 10 (*Revisar Anexo C*)

4.2.2 Pozo 2

Los resultados obtenidos para el pozo 2, considerando cada uno de los apartados, son los siguientes:

- a) **Litología.** 9 zonas (0, 1, 2, 3, 4, 5, 6, 9 y 11) del registro GR muestran lecturas por debajo de la media (112), las zonas restantes (9 y 10) muestran valores no muy lejanos a la media, pero superiores. Las respuestas del SP son todas declinaciones positivas, por lo que probablemente no sean zonas permeables.

Tabla 17. Respuesta de registros litológicos del pozo 2

ZONAS	Formación	Intervalos	Espesor(ft)	GR	SP
Zona 0	...	8650-9330	680	99,121	52,286
Zona 1	CALIZA-A	9330-9439	109	76,431	57,07
Zona 2	U SUPERIOR	9439-9492	53	82,077	51,477
Zona 3	U INFERIOR	9492-9568	76,5	70,474	28,994
Zona 4	...	9568-9674	105,5	94,091	60,912
Zona 5	CALIZA B	9674-9738	64	100,41	47,752
Zona 6	T SUPERIOR	9738-9796	58	78,495	42,897
Zona 7	T INFERIOR	9796-9830	34,5	114,55	53,156
Zona 8	...	9830-9949	118,5	122,03	62,679
Zona 9	H SUPERIOR	9949-9994	45	92,957	60,382
Zona 10	H INFERIOR	9994-10458	464	87,592	70,297

- b) **Porosidad.** El volumen de arcilla de 2 zonas (7 y 8) superan los 0.30, mientras que 5 zonas (0, 4, 5, 9 y 10) reflejan un contenido de arcilla entre 0.20-0.25 y, por último 4 zonas (1, 2, 3 y 6) muestran contenido de arcilla de 0.15-0.20.

La porosidad derivada del registro de densidad, en 2 zonas (4 y 8) muestran valores superiores a 0.4, 4 zonas (0, 2, 6 y 7) apuntan a un rango de 0.20-0.30, 2 zonas (1 y 10) se encuentran con valores inferiores a 0.1.

Tabla 18. Resultados de cálculos de porosidad del pozo 2

ZONAS	Formación	Intervalos	Espesor(ft)	VSH	DPHI
Zona 0	...	8650-9330	680	0,238	0,308
Zona 1	CALIZA-A	9330-9439	109	0,159	0,096
Zona 2	U SUPERIOR	9439-9492	53	0,174	0,246
Zona 3	U INFERIOR	9492-9568	76,5	0,143	0,141
Zona 4	...	9568-9674	105,5	0,219	0,418
Zona 5	CALIZA B	9674-9738	64	0,245	0,129
Zona 6	T SUPERIOR	9738-9796	58	0,164	0,264
Zona 7	T INFERIOR	9796-9830	34,5	0,304	0,224
Zona 8	...	9830-9949	118,5	0,332	0,379
Zona 9	H SUPERIOR	9949-9994	45	0,219	0,175
Zona 10	H INFERIOR	9994-10458	464	0,236	0,048

- c) **Resistividad.** Los valores de MNOR y MINV son bastantes cercanos, por lo que no reflejan presencia de permeabilidad en la zona.

Las respuestas más llamativas por los registros de resistividad, son en la zona invadida, para las zonas 1 y 6.

Tabla 19. Respuesta de registros resistivos del pozo 2

ZONAS	Formación	Intervalos	Espesor(ft)	MINV	MNOR	RESM	RESD
Zona 0	...	8650-9330	680	5,412	4,905	34,429	8,503
Zona 1	CALIZA-A	9330-9439	109	12,68	12,75	107,12	18,62
Zona 2	U SUPERIOR	9439-9492	53	5,29	5,168	33,865	3,697
Zona 3	U INFERIOR	9492-9568	76,5	7,835	8,377	66,667	55,62
Zona 4	...	9568-9674	105,5	5,947	5,92	25,901	5,862
Zona 5	CALIZA B	9674-9738	64	10,77	8,976	72,557	9,08
Zona 6	T SUPERIOR	9738-9796	58	7,767	6,55	31,693	6,64
Zona 7	T INFERIOR	9796-9830	34,5	11,61	11,79	118,31	17,16
Zona 8	...	9830-9949	118,5	5,168	5,142	41,256	7,564
Zona 9	H SUPERIOR	9949-9994	45	8,921	8,806	58,058	69,61
Zona 10	H INFERIOR	9994-10458	464	8,917	9,878	45,597	69,48

- d) **Presencia de hidrocarburo.** Los valores de saturación de agua que resultan más rentables son de la zona 3, 7, 9 y 10.

Tabla 20. Resultados de saturación de agua del pozo 2

ZONAS	Formación	Intervalos	Espesor(ft)	Sw	Sh
Zona 0	...	8650-9330	680	0,659	...
Zona 1	CALIZA-A	9330-9439	109	0,446	...
Zona 2	U SUPERIOR	9439-9492	53	1	...
Zona 3	U INFERIOR	9492-9568	76,5	0,258	0,742
Zona 4	...	9568-9674	105,5	0,794	...
Zona 5	CALIZA B	9674-9738	64	0,638	...
Zona 6	T SUPERIOR	9738-9796	58	0,833	...
Zona 7	T INFERIOR	9796-9830	34,5	0,464	0,536
Zona 8	...	9830-9949	118,5	0,699	...
Zona 9	H SUPERIOR	9949-9994	45	0,183	0,817
Zona 10	H INFERIOR	9994-10458	464	0,112	0,888

Tomando en cuenta lo descrito con anterioridad para el pozo 2, se determinan como prospectos a las zonas 1 y 7 (*Revisar Anexo C*)

4.2.3 Pozo 3

Los resultados obtenidos para el pozo 3, considerando cada uno de los apartados, son los siguientes:

- a) **Litología.** La respuesta del registro GR de 5 zonas se encuentra por debajo de la media (126), indicando bajo contenido de arcilla. Todas las zonas muestran una declinación negativa de SP, resaltando la zona 6 al ser la mayor, lo que indicaría presencia de permeabilidad. El registro PEF ayudo a identificar la litología de las zonas, 3 zonas (0, 2, 6) se determinan como dolomita, 2 zonas (1 y 5) como calizas, 2 zonas (3 y 10) como arenas y las 2 zonas (4 y 8) restantes como lutitas.

Tabla 21. Respuestas de registros litológicos del pozo 3

ZONAS	Formación	Intervalos	Espesor(ft)	GR	SP	PEF
Zona 0	...	8800-9946	666	102	-19,056	3,136
Zona 1	CALIZA-A	9466-9572	106	79,197	-8,4483	4,301
Zona 2	U SUPERIOR	9572-9623	51	84,822	-11,722	3,689
Zona 3	U INFERIOR	9623-9668	45	100,96	-4,5293	3,063
Zona 4	...	9668-9799	131	73,452	-3,3106	4,37
Zona 5	CALIZA B	9799-9839	40	97,871	-16,563	3,789
Zona 6	T SUPERIOR	9839-9903	64	89,226	-23,321	3,586
Zona 7	T INFERIOR	9903-9936	33	52,257	-58,553	2,269
Zona 8	...	9936-10068	132	155,14	-6,9125	3,564
Zona 9	H SUPERIOR	10068-10091	23	110,77	-10,614	3,412
Zona 10	H INFERIOR	10091-10210	119	100,3	-15,798	2,372

b) **Porosidad.** El volumen de arcilla que resalta es de la zona 8, siendo 0.42. 5 zonas (0, 3, 5, 9 y 10) se encuentran en un rango de 0.2-0.25. 4 zonas (1, 2, 4 y 6) más tienen un rango de 0.10-0.20 y solo la zona 7 presenta un valor menor a 0.1.

La porosidad total de la zona 2 es 0.3, considerada “muy buena”. 3 zonas (0, 3, 8) presentan un intervalo de 0.25-0.30, siendo “muy buena”. 3 zonas (4, 6 y 7) tienen valores de 0.20-0.25, denominándola como “muy buena”. 2 zonas (9 y 10) cuentan con respuestas de 0.15-0.20, designándola como “buena”, 2 zonas más (1 y 5) presentan valores de 0.10-0.15 calificándola como “moderada”.

La porosidad efectiva de 5 zonas (0, 2, 3, 6 y 7) presenta un resultado mayor a 0.20, 2 zonas (4 y 8) presentan valores de 0.15-0.20, 3 zonas (1, 9 y 10) cuentan con valores de 0.10-0.15 y solo la zona 5 tiene un resultado menor a 0.1.

Tabla 22. Resultados de cálculos de porosidad del pozo 3

ZONAS	Formación	Intervalos	Espesor(ft)	VSH	PHIE	TPHI
Zona 0	...	8800-9946	666	0,213	0,217	0,276
Zona 1	CALIZA-A	9466-9572	106	0,145	0,114	0,133
Zona 2	U SUPERIOR	9572-9623	51	0,16	0,26	0,31
Zona 3	U INFERIOR	9623-9668	45	0,212	0,217	0,276
Zona 4	...	9668-9799	131	0,132	0,181	0,209
Zona 5	CALIZA B	9799-9839	40	0,2	0,086	0,108
Zona 6	T SUPERIOR	9839-9903	64	0,171	0,202	0,243
Zona 7	T INFERIOR	9903-9936	33	0,079	0,204	0,222
Zona 8	...	9936-10068	132	0,418	0,157	0,269
Zona 9	H SUPERIOR	10068-10091	23	0,248	0,142	0,189
Zona 10	H INFERIOR	10091-10210	119	0,23	0,122	0,158

c) **Resistividad.** La respuesta más atractiva de los registros de resistividad son en la zona 3, 7 y 10 con el registro de resistividad profunda.

Tabla 23. Respuesta de registros resistivos del pozo 3

ZONAS	Formación	Intervalos	Espesor(ft)	MINV	MNOR	RESM	RESD
Zona 0	...	8800-9946	666	2,231	2,68	18,64	30,734
Zona 1	CALIZA-A	9466-9572	106	4,975	7,668	49,477	48,676
Zona 2	U SUPERIOR	9572-9623	51	1,982	2,983	10,792	8,1038
Zona 3	U INFERIOR	9623-9668	45	2,719	4,789	60,657	121,18
Zona 4	...	9668-9799	131	3,968	6,906	71,143	61,686
Zona 5	CALIZA B	9799-9839	40	6,033	7,442	35,091	20,51
Zona 6	T SUPERIOR	9839-9903	64	2,259	4,116	12,194	10,881
Zona 7	T INFERIOR	9903-9936	33	3,657	5,56	40,144	108,31
Zona 8	...	9936-10068	132	2,413	3,523	11,551	9,9566
Zona 9	H SUPERIOR	10068-10091	23	4,065	7,29	50,485	89,782
Zona 10	H INFERIOR	10091-10210	119	2,336	4,619	91,695	256,14

d) **Presencia de hidrocarburo.** Las zonas 3, 7, 9 y 10 cuentan con valores de saturación de agua menores a 0.30, lo que las convierte en sugerentes.

Tabla 24. Resultados de saturación de agua del pozo 3

ZONAS	Formación	Intervalos	Espesor(ft)	Sw	Sh
Zona 0	...	8800-9946	666	0,513	...
Zona 1	CALIZA-A	9466-9572	106	0,408	...
Zona 2	U SUPERIOR	9572-9623	51	1	...
Zona 3	U INFERIOR	9623-9668	45	0,259	0,741
Zona 4	...	9668-9799	131	0,362	...
Zona 5	CALIZA B	9799-9839	40	0,629	...
Zona 6	T SUPERIOR	9839-9903	64	0,918	...
Zona 7	T INFERIOR	9903-9936	33	0,274	0,726
Zona 8	...	9936-10068	132	0,902	...
Zona 9	H SUPERIOR	10068-10091	23	0,217	0,726
Zona 10	H INFERIOR	10091-10210	119	0,123	0,877

Con lo relatado con anterioridad, las zonas prospecto para el pozo 3, son la zona 3, 7 y 10 (Revisar anexo C)

4.2.4 Pozo 4

Los resultados obtenidos para el pozo 4, considerando cada uno de los apartados, son los siguientes:

a) **Litología.** Las respuestas de 9 zonas (0, 1, 2, 3, 4, 5, 6, 7 y 10) del registro GR son bajas, la zona 8 presenta una lectura de 104.23, superando la media (74.94) del registro en el pozo y la zona 9 presenta una lectura por debajo de la media, pero más alta que las 9 zonas anteriores. El registro SP, muestra respuestas positivas en todas las zonas.

Tabla 25. Respuesta de registros litológicos del pozo 4

ZONAS	Formación	Intervalos	Espesor(ft)	GR	SP
Zona 0	...	8700-9370	670	67,114	73,03
Zona 1	CALIZA-A	9370-9477	107	50,96	75,598
Zona 2	U SUPERIOR	9477-9537	60	61,906	70,071
Zona 3	U INFERIOR	9537-9580	43	61,384	48,897
Zona 4	...	9580-9714	134	46,232	79,895
Zona 5	CALIZA B	9714-9770	56	74,131	66,501
Zona 6	T SUPERIOR	9770-9381	61	56,454	53,586
Zona 7	T INFERIOR	9381-9879	48	33,837	36,541
Zona 8	...	9879-9997	118	104,23	78,508
Zona 9	H SUPERIOR	9997-10032	35	82,486	68,375
Zona 10	H INFERIOR	10032-10152	120	42,187	85,373

b) **Porosidad.** 3 zonas (5, 8 y 9) reflejan un volumen de arcilla por encima de 0.3, resaltando la zona 8, con 0.55. 4 zonas (0, 2, 3 y 6) reflejan valores en un rango de 0.20-0.30. Y, por último 4 zonas (1, 4, 7 y 10) muestran que lo contienen en un valor menor a 0.20.

La porosidad total, derivada del registro de densidad es mayor a 0.20 en 6 zonas (0, 2, 3, 6, 7 y 8) de este pozo, 2 zonas (4 y 9) se encuentran en el rango de 0.15-0.20. La zona 10 tiene un valor de 0.13 y la zona 1 cuenta con un valor menor a 0.1

Tabla 26. Resultados de cálculos de porosidad del pozo 4

ZONAS	Formación	Intervalos	Espesor(ft)	VSH	DPHI
Zona 0	...	8700-9370	670	0,284	0,249
Zona 1	CALIZA-A	9370-9477	107	0,199	0,092
Zona 2	U SUPERIOR	9477-9537	60	0,261	0,269
Zona 3	U INFERIOR	9537-9580	43	0,277	0,21
Zona 4	...	9580-9714	134	0,186	0,159
Zona 5	CALIZA B	9714-9770	56	0,331	0,086
Zona 6	T SUPERIOR	9770-9381	61	0,227	0,211
Zona 7	T INFERIOR	9381-9879	48	0,125	0,257
Zona 8	...	9879-9997	118	0,548	0,266
Zona 9	H SUPERIOR	9997-10032	35	0,397	0,152
Zona 10	H INFERIOR	10032-10152	120	0,176	0,129

c) **Resistividad.** Los valores de MNOR son menores a MINV, lo que indica zonas no permeables en la formación. La respuesta del registro de profundidad resalta para las zonas 4, 7 y 10.

Tabla 27. Respuestas de registros resistivos del pozo 4

ZONAS	Formación	Intervalos	Espesor(ft)	MINV	MNOR	RESM	RESD
Zona 0	...	8700-9370	670	4,6805	3,69	37,594	8,193
Zona 1	CALIZA-A	9370-9477	107	12,591	11,39	59,907	22,75
Zona 2	U SUPERIOR	9477-9537	60	5,6763	4,235	24,867	5,294
Zona 3	U INFERIOR	9537-9580	43	3,2062	2,661	35,631	61,86
Zona 4	...	9580-9714	134	14,689	14,14	77,287	104,3
Zona 5	CALIZA B	9714-9770	56	10,748	8,705	20,47	11,95
Zona 6	T SUPERIOR	9770-9381	61	6,886	5,256	14,508	7,45
Zona 7	T INFERIOR	9381-9879	48	3,2801	2,991	77,998	127,1
Zona 8	...	9879-9997	118	4,4249	3,603	12,742	6,949
Zona 9	H SUPERIOR	9997-10032	35	9,101	8,453	91,831	29,14
Zona 10	H INFERIOR	10032-10152	120	5,1756	5,479	131,16	122,8

d) **Presencia de hidrocarburo.** La zona 3, 4, 7 y 10 presentan buenas respuestas de saturación de agua. Por lo que, se estima la posible saturación de hidrocarburo de las zonas.

Tabla 28. Resultados de saturación de agua del pozo 4

ZONAS	Formación	Intervalos	Espesor(ft)	Sw	Sh
Zona 0	...	8700-9370	670	0,804	...
Zona 1	CALIZA-A	9370-9477	107	0,482	...
Zona 2	U SUPERIOR	9477-9537	60	1	...
Zona 3	U INFERIOR	9537-9580	43	0,293	0,707
Zona 4	...	9580-9714	134	0,225	0,775
Zona 5	CALIZA B	9714-9770	56	0,666	...
Zona 6	T SUPERIOR	9770-9381	61	0,739	...
Zona 7	T INFERIOR	9381-9879	48	0,204	0,796
Zona 8	...	9879-9997	118	0,873	...
Zona 9	H SUPERIOR	9997-10032	35	0,291	...
Zona 10	H INFERIOR	10032-10152	120	0,151	0,849

En cuanto a lo relatado, se identifican como zonas prospecto la zona 4, 7 y 10 para este pozo (*Revisar Anexo C*)

4.2.5 Pozo 5

Los resultados obtenidos para el pozo 5, considerando cada uno de los apartados, son los siguientes:

- a) **Litología.** La respuesta de 8 zonas (1, 2, 3, 4, 5, 6, 7 y 10) presentan lecturas de Gr por debajo de la media (93.64), 2 zonas (0 y 9) presentan valores superiores, pero cercanos a la media y la zona 8 presenta un valor relativamente alto, considerando que es una formación con alto contenido de arcilla.

Las respuestas del registro SP para todas las zonas son positivas.

Tabla 29. Respuesta de registros litológicos del pozo 5

ZONAS	Formación	Intervalos	Espesor(ft)	GR	SP
Zona 0	...	8750-9381	631	93,687	83,552
Zona 1	CALIZA-A	9381-9494	113	70,418	83,237
Zona 2	U SUPERIOR	9494-9560	66	84,125	85,153
Zona 3	U INFERIOR	9560-9586	26	46,918	22,026
Zona 4	...	9586-9724	138	50,358	82,803
Zona 5	CALIZA B	9724-9785	61	82,114	43,997
Zona 6	T SUPERIOR	9785-9838	53	75,752	28,961
Zona 7	T INFERIOR	9838-9898	60	49,137	-6,363
Zona 8	...	9898-10016	118	134,5	49,927
Zona 9	H SUPERIOR	10016-10085	69	94,302	50,777
Zona 10	H INFERIOR	10085-10130	45	34,019	64,935

- b) **Porosidad.** El volumen de arcilla de 3 zonas (0, 8 y 9) supera los 0.3, resaltando la zona 8, con 0.567. Los valores de 3 zonas (2, 5 y 6) son mayores a 0.20 y menores a 0.30. La cantidad de 4 zonas (1, 3, 4 y 7) son valores entre 0.10-0.20. Para finalizar, la zona 10 cuenta con un monto menor a 0.1.

La porosidad derivada del registro de densidad es mayor a 0.3 en 2 zonas (2 y 10). En 4 zonas (0, 6, 7 y 8) se encuentra en un rango de 0.20-0.25. En 3 zonas (3, 4 y 9) está en un rango de 0.15-0.20. Finalmente, la zona 5 tiene un valor menor a 0.1.

Tabla 30. Resultados de cálculos de porosidad del pozo 5

ZONAS	Formación	Intervalos	Espesor(ft)	VSH	DPHI
Zona 0	...	8750-9381	631	0,296	0,249
Zona 1	CALIZA-A	9381-9494	113	0,19	0,124
Zona 2	U SUPERIOR	9494-9560	66	0,257	0,382
Zona 3	U INFERIOR	9560-9586	26	0,112	0,206
Zona 4	...	9586-9724	138	0,119	0,172
Zona 5	CALIZA B	9724-9785	61	0,239	0,073
Zona 6	T SUPERIOR	9785-9838	53	0,21	0,201
Zona 7	T INFERIOR	9838-9898	60	0,107	0,241
Zona 8	...	9898-10016	118	0,567	0,209
Zona 9	H SUPERIOR	10016-10085	69	0,305	0,179
Zona 10	H INFERIOR	10085-10130	45	0,052	0,339

c) **Resistividad.** Los valores de resistividad de MNOR y MINV son bastantes cercanos, por lo que no indican presencia de permeabilidad. La respuesta de RESM para la zona 2 y 5 es sugerente.

Tabla 31. Respuesta de registros resistivos del pozo 5

ZONAS	Formación	Intervalos	Espesor(ft)	MINV	MNOR	RESM	RESD
Zona 0	...	8750-9381	631	2,813	2,838	54,687	5,1166
Zona 1	CALIZA-A	9381-9494	113	7,734	7,484	41,16	15,693
Zona 2	U SUPERIOR	9494-9560	66	3,093	2,764	115,27	3,3985
Zona 3	U INFERIOR	9560-9586	26	2,217	3,381	1,7489	62,763
Zona 4	...	9586-9724	138	7,926	7,853	36,892	62,758
Zona 5	CALIZA B	9724-9785	61	8,388	7,744	155,76	12,973
Zona 6	T SUPERIOR	9785-9838	53	5,204	5,074	1,7823	6,3263
Zona 7	T INFERIOR	9838-9898	60	2,999	4,135	2,4955	46,93
Zona 8	...	9898-10016	118	3,838	3,479	39,959	5,0343
Zona 9	H SUPERIOR	10016-10085	69	6,356	7,049	4,9476	53,623
Zona 10	H INFERIOR	10085-10130	45	1,629	2,592	1,4224	57,356

d) **Presencia de hidrocarburo.** Las mejores respuestas de saturación de agua se presentan en las zonas 3, 4, 7, 9 y 10.

Tabla 32. Resultados de saturación de agua del pozo 5

ZONAS	Formación	Intervalos	Espesor(ft)	Sw	Sh
Zona 0	...	8750-9381	631	0,815	...
Zona 1	CALIZA-A	9381-9494	113	0,465	...
Zona 2	U SUPERIOR	9494-9560	66	1	...
Zona 3	U INFERIOR	9560-9586	26	0,233	0,767
Zona 4	...	9586-9724	138	0,233	0,767

Zona 5	CALIZA B	9724-9785	61	0,512	...
Zona 6	T SUPERIOR	9785-9838	53	0,649	...
Zona 7	T INFERIOR	9838-9898	60	0,269	0,731
Zona 8	...	9898-10016	118	0,822	...
Zona 9	H SUPERIOR	10016-10085	69	0,152	0, 848
Zona 10	H INFERIOR	10085-10130	45	0,339	0,661

4.3 Evaluación de reservas probables

Tabla 33. Cálculos de las reservas probables

ZONAS	Formaciones	Intervalos	Espesor(ft)	Reservas posibles	
Zona 0	...	8800-9946	666		
Zona 1	CALIZA-A	9466-9572	106		
Zona 2	U SUPERIOR	9572-9623	51		
Zona 3	U INFERIOR	9623-9668	45	22.874.307.734.813	22 MMM BY
Zona 4	...	9668-9799	131		
Zona 5	CALIZA B	9799-9839	40		
Zona 6	T SUPERIOR	9839-9903	64		
Zona 7	T INFERIOR	9903-9936	33	13.910.689.429.052	13 MMBY
Zona 8	...	9936-10068	132		
Zona 9	H SUPERIOR	10068-10091	23		
Zona 10	H INFERIOR	10091-10210	119	35.373.497.224.722	35 MMBY

En los resultados obtenidos del cálculo de las posibles reservas a partir del análisis realizado anteriormente con los registros de pozo, se obtiene que la formación podría contener 22 MMBY para U inferior, 13 MMBY para T inferior y 35 MMBY para H inferior.

CAPÍTULO V: CONCLUSIONES Y RECOMENDACIONES

5.1 Conclusiones

- Se recolectó la información de 5 pozos de los 20 del campo, a utilizar en el código desarrollado en este trabajo, usando un muestreo por conveniencia; tomando en cuenta la mayor cantidad de registros de pozo con la que contaban cada uno de ellos.
- Se logró organizar de manera exitosa la información en una hoja de cálculo de Excel, bajo el uso del código generado, reconociendo zonas con datos faltantes y rellenando su información bajo la media del valor para la formación específica.
- Se distribuyó los datos en un conjunto 80/20, siendo 20% para entrenamiento y 80% para prueba, resultando en una media del 11% de diferencia entre los resultados de coeficiente de determinación (R^2) para prueba y entrenamiento, respectivamente.
- Se determinó el modelo lineal, polinomial, máquina de soporte vectorial y árboles de regresión de los modelos de regresión de ML disponibles para medir la calidad de los datos. Llevando en aumento la complejidad partiendo del más simple como es el modelo lineal, al de mayor complejidad como es el de árboles de regresión.
- Se mejoró los modelos que obtuvieron una métrica de error menor al 0.7, siendo el mejor resultado el de gradient boosting regressor, por tener un resultado de 0.8. También, se probó Bosques de regresión, pero este presentó un sobreajuste.

- Se determinaron las profundidades de las formaciones, con la utilización de los registros disponibles, denominándolos por zonas, desde la 0 a la 10. Con un máximo de profundidad alrededor de 10000 ft.
- Las medidas de las propiedades petrofísicas resultaron ser más atractivas en la zona 3 y 10 para el pozo 1. En el pozo 2, fueron más atractivas la zona 1 y 7. En el pozo 3, lo fueron las zonas 3, 7 y 10. Para el pozo 4, resultaron ser las zonas 4, 7 y 10. Y, finalmente en el pozo 5, lo fueron la zona 2 y 5. Con todo esto detallado, tomando en cuenta a los 5 pozos y la información disponible, se tomó en cuenta las zonas 3, 7 y 10 para el cálculo de las reservas probables.
- Determinando las reservas para la formación U inferior, T superior, H inferior, siendo estas la zona 3, 7 y 10. Al ser el pozo 1 y 3, los únicos que cuentan con el cálculo de porosidad total (TPHI), se consideran los resultados de TPHI del pozo 3 para el cálculo, se obtuvo 22MMMBY para U inferior, 13 MMBY para T inferior y 35 MMBY para H inferior.
- Se analizaron los resultados obtenidos por el algoritmo en contraste con los que presentan los respectivos registros, y teniendo en cuenta que solo se cuenta con PHIE, VSH y Sw para realizar las comparaciones. De este análisis se obtuvieron los siguientes resultados: en el pozo 1, discrepan en un 50% el cálculo de volumen de arcilla y la porosidad efectiva, mientras que la saturación de agua, en un 20%; en el pozo 2, se diferencian en un 70% los valores de volumen de arcilla y un 0.40 en la saturación de agua; en el pozo 3, se diferencian en un 60% en el cálculo de volumen de arcilla, 70% en la porosidad efectiva y 27% en la saturación de agua; en el pozo 4, presentan discordancia en un 36% en volumen de arcilla y 37% en saturación de agua; en el pozo 5, se distinguen por el 58% en el volumen de arcilla y 36% en la saturación de agua.

5.2 Recomendaciones

- Usar más pozos con el fin de caracterizar el comportamiento de los datos por registro de pozo, pues 5 pozos no son suficientes para definir el comportamiento de los datos, más aún cuando hay bastante variabilidad en los datos de estos 5 pozos.
- De ser posible y contar con la información suficiente, usar mejores ecuaciones para el cálculo de las reservas probables de las formaciones de los pozos.
- De ser posible, tomar en cuenta la información de pruebas de laboratorio y análisis de núcleo.
- Probar el uso de Deep learning en el análisis de los datos, al no obtener resultados validos con ML.

REFERENCIAS BIBLIOGRÁFICAS

- Adesina, E., & Olusola, B. (2022). Application of Machine Learning Algorithm for Predicting Produced Water Under Various Operating Conditions in an Oilwell . *One Petro.*
- Ahmed, S. A., MonaLisa, Hussain, M., & Khan, Z. U. (2022). Supervised machine learning for predicting shear sonic log (DTS) and volumes of petrophysical and elastic attributes, Kadanwari Gas Field, Pakistan. *Frontiers in Earth Science*, 10. <https://doi.org/10.3389/feart.2022.919130>
- Asquith, G., & Gibson, C. (1982). *Basic Well Log Analysis For Geologists.*
- Balz, O., Pivot, F., & Veeken, P. (2015, January 7). *Reservoir Characterisation Using Neural Networks Controlled by Petrophysical and Seismic Modelling.* <https://doi.org/10.3997/2214-4609.201407673>
- Belyadi, H., & Haghghat, A. (2021). *Machine Learning Guide for Oil and Gas Using Python.* Elsevier. <https://doi.org/10.1016/b978-0-12-821929-4.01001-5>
- Bittar, M., Wang, S., Wu, X., & Chen, J. (2021). Multiple Well-Log Depth Matching Using Deep Q-Learning. *Petrophysics.*
- Camargo, J. (2005). *Introducción a la interpretación de perfiles de pozo abierto.*
- Dandekar, A. Y. (2013). *Petroleum reservoir rock and fluid properties.*
- Escobar, F. H. (2012). *Fundamentos de ingeniería de yacimientos.*
- Fanchi, J., & Christiansen, R. (2017). *Introducción to Petroleum engineering.*
- Garduza, V. M. (2019). *Manual de Registros Geofísicos de Pozo y algunas Aplicaciones.*
- Gi, M., Gokaraju, D., Chen, D., & Quiren, J. (2016). Shale Fracturing Characterization and Optimization by Using Anisotropic Acoustic Interpretation, 3D Fracture Modeling, and Supervised Machine Learning. *Petrophysics.*

- Grajales, T. (n.d.). Tipos de investigación. Retrieved July 9, 2023, from <https://cmapspublic2.ihmc.us/rid=1RM1F0L42-VZ46F4-319H/871.pdf>
- Gu, Y., Gao, Z., & Rui, Z. (2018). Complex Lithofacies Identification Using Improved Probabilistic Neural Networks. *Petrophysics*.
- Holstein, E. D. (2007). *Petroleum Engineering Handbook*.
- Kamel, M. H., & Mabrouk, W. M. (2003). Estimation of shale volume using a combination of the three porosity logs. *Journal of Petroleum Science and Engineering*, 40(3–4), 145–157. [https://doi.org/10.1016/S0920-4105\(03\)00120-7](https://doi.org/10.1016/S0920-4105(03)00120-7)
- Lefranc, M., Bayraktar, Z., Kristensen, M., Driss, H., Le Nir, I., & Marza, P. (2021). Deep Learning Based Automated Sedimentary Geometry Characterization From Borehole Images. *Petrophysics*.
- Müller, A. C., & Guido, S. (2017). *Introduction to Machine Learning with Python*.
- Pandey, Y. N., Rastogi, A., Kainkaryam, S., Bhattacharya, S., & Saputelli, L. (2020). Machine Learning in the Oil and Gas Industry. In *Machine Learning in the Oil and Gas Industry*. Apress. <https://doi.org/10.1007/978-1-4842-6094-4>
- Paris de Ferrer, M. (2009). *Fundamentos de Ingeniería de Yacimientos*.
- Priezzhev, I., Veecken, P. C. H., Egorov, S. V., & Strecker, U. (2019). Direct prediction of petrophysical and petroelastic reservoir properties from seismic and well-log data using nonlinear machine learning algorithms. *Geophysics*. <https://doi.org/10.1190/tle38120949.1>
- Rodríguez, E., & Carecedo, J. (2022). *Modelos predictivos con Machine Learning*.
- Rogulina, A., Zaytsev, A., Ismailova, L., Kovalev, D., Katterbauer, K., & Marsala, A. (2022, February 21). *Similarity Learning for Well Logs Prediction Using Machine Learning Algorithms*. <https://doi.org/10.2523/iptc-22067-ms>
- Serra, O. (1984). *Fundamentals of well log interpretation*.
- Simoës, V., Maniar, H., Abubakar, A., & Zhao, T. (2022). Deep Learning for Multiwell Automatic Log Correction. *Petrophysics*.
- Tiab, D., & Donaldson, E. C. (2016). *Petrophysics : theory and practice of measuring reservoir rock and fluid transport properties*.

Van den Bossche, J., Fan, T., & Gramfort, A. (n.d.). *scikit-learn*. Choosing the Right Stimator.

Xu, C., Fu, L., Lin, T., Li, W., & Ma, S. (2022). Machine learning in petrophysics: Advantages and limitations. *Artificial Intelligence in Geosciences*, 3, 157–161. <https://doi.org/10.1016/j.aiig.2022.11.004>

Yang, S. (2017). *Fundamentals of Petrophysics*. <http://www.springer.com/series/10173>

ANEXOS

ANEXO A



Figura 21. Mapa de calor mostrando correlación de los registros del pozo 1



Figura 22. Mapa de calor mostrando correlación de los registros del pozo 2

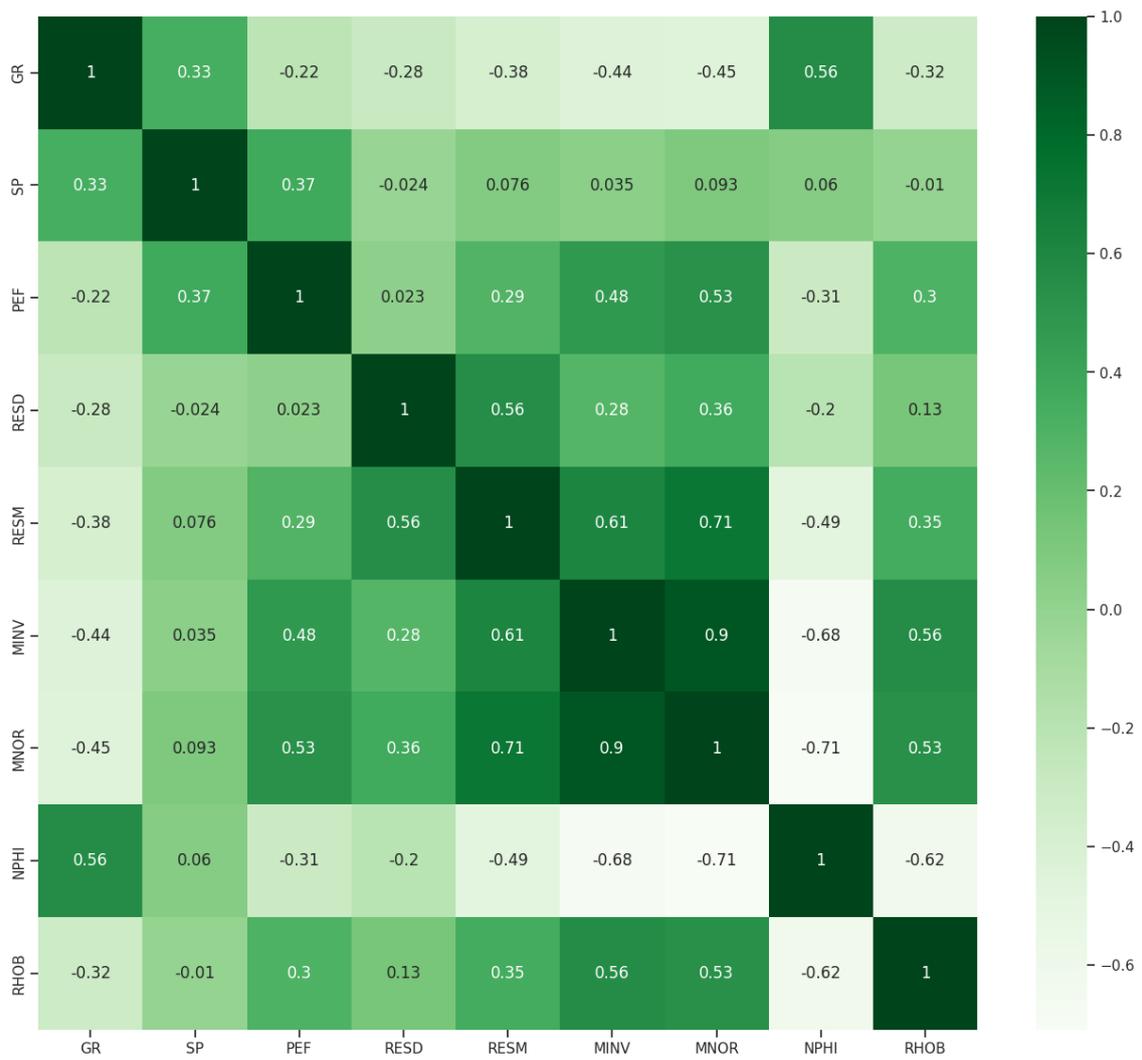


Figura 23. Mapa de calor mostrando correlación de los registros del pozo 3



Figura 24. Mapa de calor mostrando correlación de los registros del pozo 4

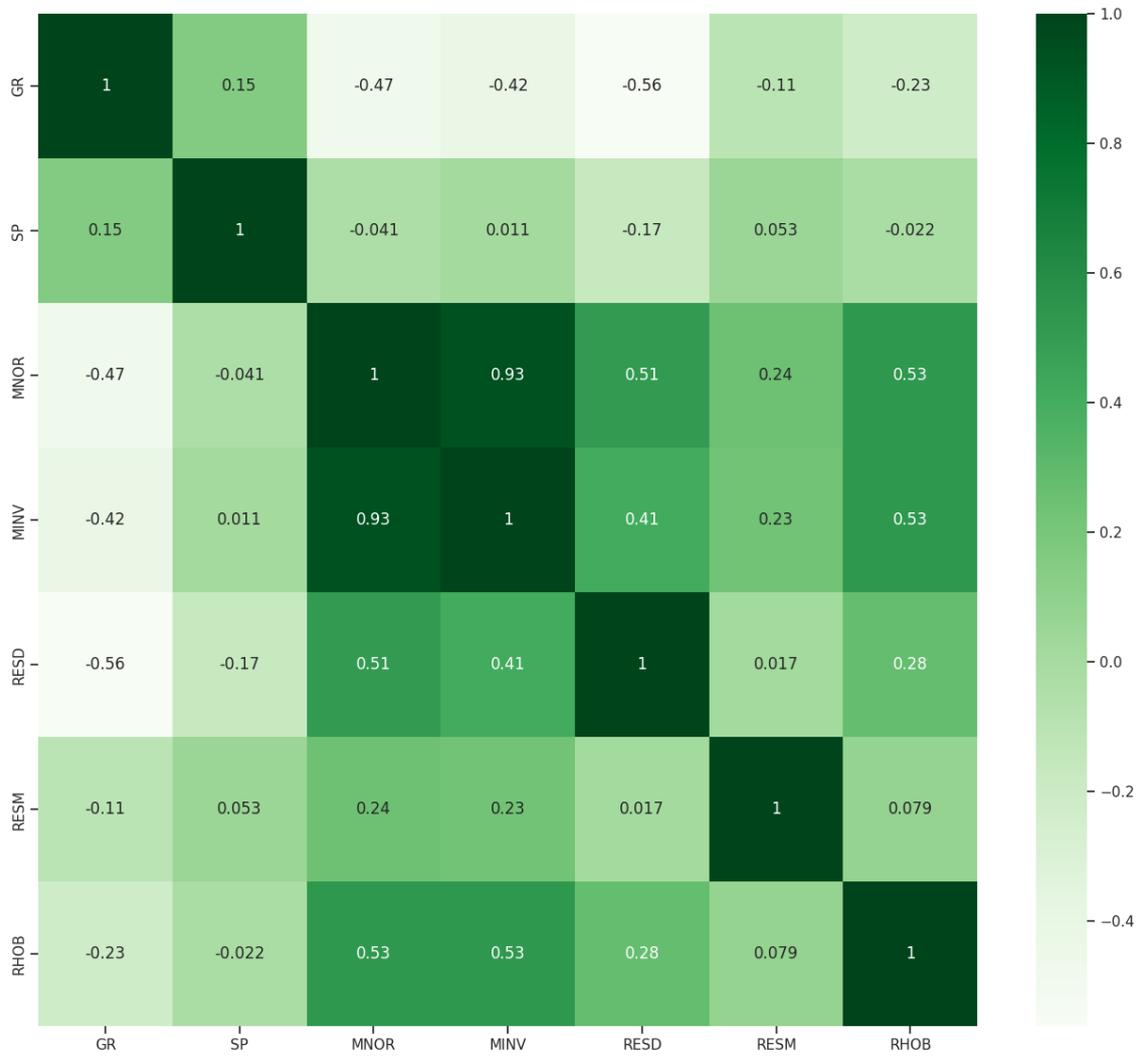


Figura 25. Mapa de calor mostrando correlación de los registros del pozo 5

ANEXO B

COMPORTAMIENTO DE LOS DATOS DE ÁRBOL DE REGRESIÓN DE PROFUNDIDAD 5:

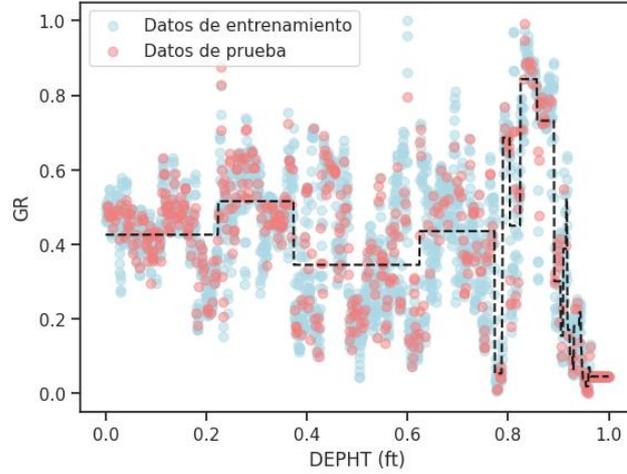


Figura 26. Modelamiento de datos para registro GR con ML

Entrenamiento GR ÁRBOL DE REGRESIÓN DE PROFUNDIDAD 5: 0.6119944884767161

Prueba GR ÁRBOL DE REGRESIÓN DE PROFUNDIDAD 5: 0.6509810711283175

COMPORTAMIENTO DE LOS DATOS DE ÁRBOL DE REGRESIÓN DE PROFUNDIDAD 5:

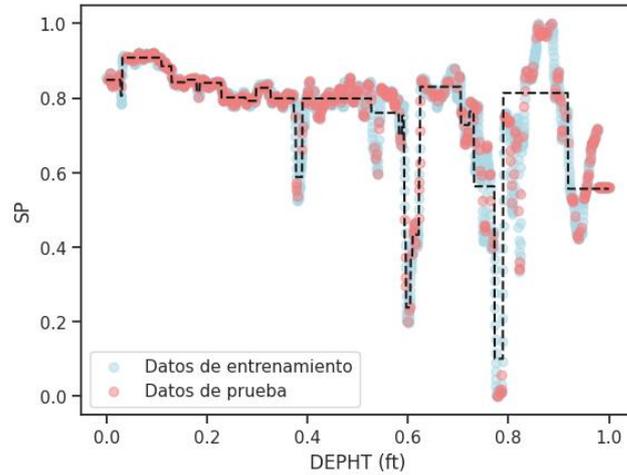


Figura 27. Modelamiento de datos para registro SP con ML

Entrenamiento SP ÁRBOL DE REGRESIÓN DE PROFUNDIDAD 5: 0.8286453693367267

Prueba SP ÁRBOL DE REGRESIÓN DE PROFUNDIDAD 5: 0.8032387346440273

COMPORTAMIENTO DE LOS DATOS CON ÁRBOL DE REGRESIÓN DE PROFUNDIDAD 3:

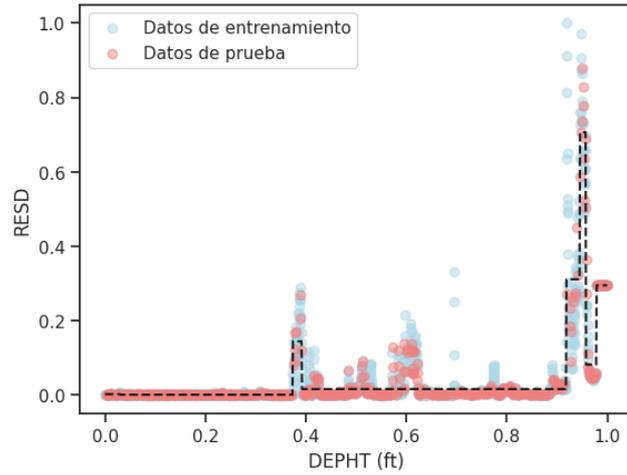


Figura 28. Modelamiento de datos para registro RESD con ML

Entrenamiento RESD ÁRBOL DE REGRESIÓN DE PROFUNDIDAD 3: 0.8387011079184712

Prueba RESD ÁRBOL DE REGRESIÓN DE PROFUNDIDAD 3: 0.9022244381079066

COMPORTAMIENTO DE LOS DATOS CON ARBOL DE REGRESIÓN DE PROFUNDIDAD 2

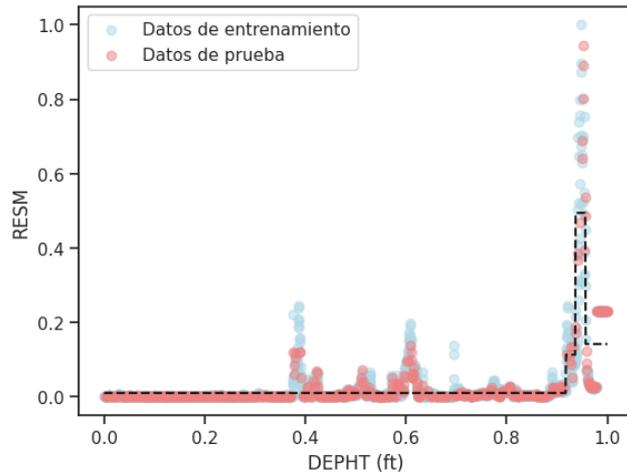


Figura 29. Modelamiento de datos para registro RESM con MI

Entrenamiento RESM ÁRBOL DE REGRESIÓN DE PROFUNDIDAD 2: 0.714437094188112

Prueba RESM ÁRBOL DE REGRESIÓN DE PROFUNDIDAD 2: 0.7513297550252631

COMPORTAMIENTO DE LOS DATOS DE ÁRBOL DE REGRESIÓN DE PROFUNDIDAD 5:

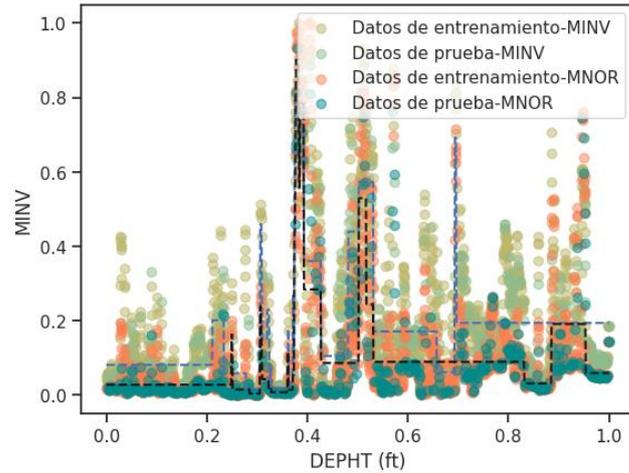


Figura 30. Modelamiento de datos para registro MNOR y MINV con ML

Entrenamiento-MINV MINV ÁRBOL DE REGRESIÓN DE PROFUNDIDAD 5: 0.5905557322519801

Prueba-MINV MINV ÁRBOL DE REGRESIÓN DE PROFUNDIDAD 5: 0.5387848902664252

Entrenamiento-MNOR MINV ÁRBOL DE REGRESIÓN DE PROFUNDIDAD 5: 0.6071424607455047

Prueba-MNOR MINV ÁRBOL DE REGRESIÓN DE PROFUNDIDAD 5: 0.5595906977428481

COMPORTAMIENTO DE LOS DATOS DE ÁRBOL DE REGRESIÓN DE PROFUNDIDAD 5:

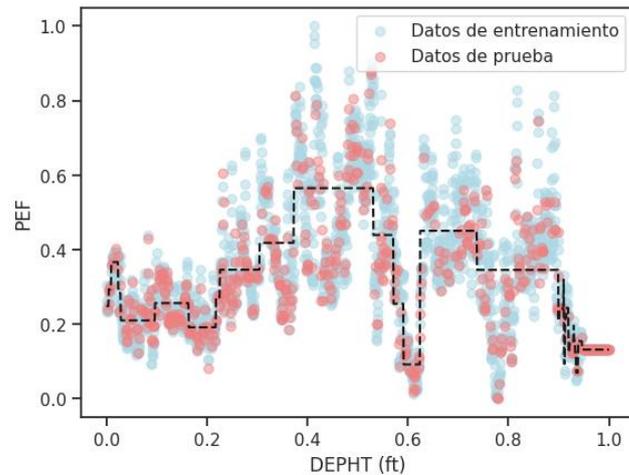


Figura 31. Modelamiento de datos para registro PEF con ML

Entrenamiento PEF ÁRBOL DE REGRESIÓN DE PROFUNDIDAD 5: 0.595779962185251

Prueba PEF ÁRBOL DE REGRESIÓN DE PROFUNDIDAD 5: 0.5974312871042966

COMPORTAMIENTO DE LOS DATOS DE ÁRBOL DE REGRESIÓN DE PROFUNDIDAD 5:

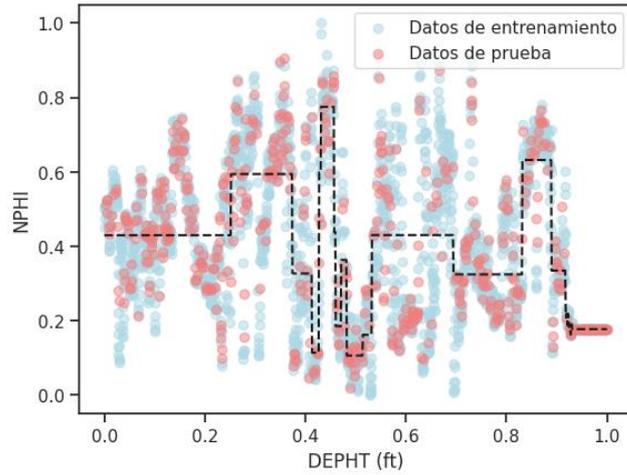


Figura 32. Modelamiento de datos para registro NPHI con ML

Entrenamiento NPHI ÁRBOL DE REGRESIÓN DE PROFUNDIDAD 5: 0.5200125754499094

Prueba NPHI ÁRBOL DE REGRESIÓN DE PROFUNDIDAD 5: 0.5405903847120748

COMPORTAMIENTO DE LOS DATOS DE ÁRBOL DE REGRESIÓN DE PROFUNDIDAD 5:

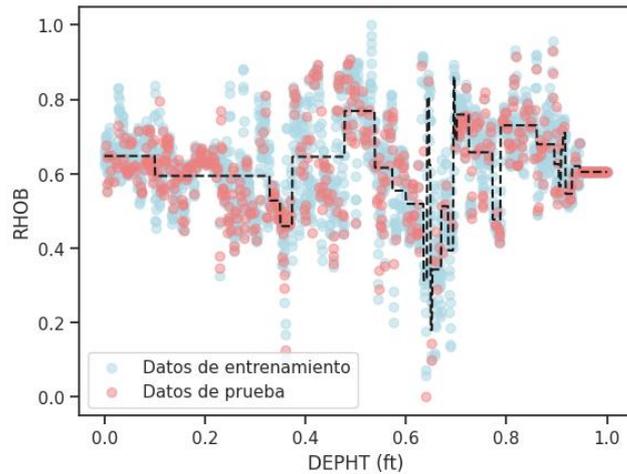


Figura 33. Modelamiento de datos para registro RHOB con ML

Entrenamiento RHOB ÁRBOL DE REGRESIÓN DE PROFUNDIDAD 5: 0.530740389634688

Prueba RHOB ÁRBOL DE REGRESIÓN DE PROFUNDIDAD 5: 0.4525454447615238

ANEXO C

Tabla 34. Resultados de evaluación petrofísica del pozo 1

Zonas	Formación	Intervalos	Espesor(ft)	REGISTROS							CÁLCULOS PETROFÍSICOS						
				GR	SP	PEF	MINV	MNOR	RESM	RESD	VSH	PHIE	TPHI	Rw	F	Sw	Sh
Zona 0	...	8750-9385	635	95,941	-14,765	3,257	2,039	3,435	9,745	11,657	0,297	0,181	0,257	0,437	10,058	0,614	...
Zona 1	CALIZA-A	9385-9489	104	72,383	-16,932	4,523	5,241	10,27	20,59	23,94	0,185	0,138	0,17	0,252	17,457	0,429	...
Zona 2	U SUPERIOR	9489-9571	82	80,102	-20,54	3,467	2,519	4,197	11,56	14,434	0,219	0,195	0,249	0,421	10,457	0,552	...
Zona 3	U INFERIOR	9571-9617.5	46,5	70,235	-52,806	2,184	2,38	3,914	101,4	116,35	0,194	0,133	0,165	0,243	18,101	0,194	0,806
Zona 4	...	9617.5-9715	97,5	101,39	-14,645	3,699	2,132	2,674	5,512	4,3984	0,327	0,201	0,298	0,532	8,2616	1	...
Zona 5	CALIZA B	9715-9756	41	92,849	-20,095	3,935	4,178	6,105	17,07	20,815	0,28	0,138	0,191	0,295	14,887	0,46	...
Zona 6	T SUPERIOR	9756-9810	54	77,336	-31,732	3,561	2,26	3,074	8,661	7,3474	0,204	0,179	0,225	0,367	11,994	0,974	...
Zona 7	T INFERIOR	9810-9845	35	59,233	-70,889	2,267	2,465	5,718	21,81	22,593	0,143	0,136	0,159	0,232	18,999	0,441	...
Zona 8	...	9845.5-9986	141	135,44	-16,668	3,535	2,782	3,581	10,69	8,442	0,562	0,105	0,239	0,397	11,073	0,722	...
Zona 9	H SUPERIOR	9986-10025	39	80,214	-18,195	3,152	3,323	5,635	33,91	37,738	0,229	0,157	0,203	0,321	13,723	0,298	...
Zona10	H INFERIOR	10025-10140	115	36,138	-38,271	2,371	2,705	6,105	308,1	391,69	0,043	0,125	0,13	0,177	24,847	0,09	0,91

Tabla 35. Resultados de evaluación petrofísica del pozo 2

Zonas	Formación	Intervalos	Espesor(ft)	REGISTROS						CÁLCULOS PETRFÍSICOS					
				GR	SP	MINV	MNOR	RESM	RESD	VSH	DPHI	Rw	F	Sw	Sh
Zona 0	...	8650-9330	680	99,121	52,286	5,412	4,905	34,429	8,503	0,238	0,308	0,467	7,916	0,659	...
Zona 1	CALIZA-A	9330-9439	109	76,431	57,07	12,68	12,75	107,12	18,62	0,159	0,096	0,1	37,02	0,446	...
Zona 2	U SUPERIOR	9439-9492	53	82,077	51,477	5,29	5,168	33,865	3,697	0,174	0,246	0,347	10,64	1	...
Zona 3	U INFERIOR	9492-9568	76,5	70,474	28,994	7,835	8,377	66,667	55,62	0,143	0,141	0,165	22,35	0,258	0,742
Zona 4	...	9568-9674	105,5	94,091	60,912	5,947	5,92	25,901	5,862	0,219	0,418	0,702	5,267	0,794	...
Zona 5	CALIZA B	9674-9738	64	100,41	47,752	10,77	8,976	72,557	9,08	0,245	0,129	0,147	25,2	0,638	...
Zona 6	T SUPERIOR	9738-9796	58	78,495	42,897	7,767	6,55	31,693	6,64	0,164	0,264	0,381	9,699	0,833	...
Zona 7	T INFERIOR	9796-9830	34,5	114,55	53,156	11,61	11,79	118,31	17,16	0,304	0,224	0,305	12,1	0,464	0,536
Zona 8	...	9830-9949	118,5	122,03	62,679	5,168	5,142	41,256	7,564	0,332	0,379	0,617	5,99	0,699	...
Zona 9	H SUPERIOR	9949-9994	45	92,957	60,382	8,921	8,806	58,058	69,61	0,219	0,175	0,22	16,8	0,183	0,817
Zona 10	H INFERIOR	9994-10458	464	87,592	70,297	8,917	9,878	45,597	69,48	0,236	0,048	0,039	94,57	0,112	0,888

Tabla 36. Resultados de evaluación petrofísica del pozo 3

Zonas	Formación	Intervalos	Espesor(ft)	REGISTROS								CÁLCULOS PETROFÍSICOS					
				GR	SP	PEF	MIN	MNOR	RESM	RESD	VSH	PHIE	TPHI	Rw	F	Sw	Sh
Zona 0	...	8800-9946	666	101,999	-19,056	3,136	2,231	2,68	18,64	30,734	0,213	0,217	0,276	0,886	9,145	0,513	...
Zona 1	CALIZA-A	9466-9572	106	79,1969	-8,4483	4,301	4,975	7,668	49,477	48,676	0,145	0,114	0,133	0,335	24,18	0,408	...
Zona 2	U SUPERIOR	9572-9623	51	84,8221	-11,722	3,689	1,982	2,983	10,792	8,1038	0,16	0,26	0,31	1,033	7,848	1	...
Zona 3	U INFERIOR	9623-9668	45	100,96	-4,5293	3,063	2,719	4,789	60,657	121,18	0,212	0,217	0,276	0,886	9,143	0,259	0,741
Zona 4	...	9668-9799	131	73,4517	-3,3106	4,37	3,968	6,906	71,143	61,686	0,132	0,181	0,209	0,613	13,22	0,362	...
Zona 5	CALIZA B	9799-9839	40	97,8715	-16,563	3,789	6,033	7,442	35,091	20,51	0,2	0,086	0,108	0,255	31,82	0,629	...
Zona 6	T SUPERIOR	9839-9903	64	89,2257	-23,321	3,586	2,259	4,116	12,194	10,881	0,171	0,202	0,243	0,75	10,81	0,918	...
Zona 7	T INFERIOR	9903-9936	33	52,2574	-58,553	2,269	3,657	5,56	40,144	108,31	0,079	0,204	0,222	0,663	12,22	0,274	0,726
Zona 8	...	9936-10068	132	155,14	-6,9125	3,564	2,413	3,523	11,551	9,9566	0,418	0,157	0,269	0,857	9,459	0,902	...
Zona 9	H SUPERIOR	10068- 10091	23	110,767	-10,614	3,412	4,065	7,29	50,485	89,782	0,248		0,189	0,536	15,11	0,217	...
Zona 10	H INFERIOR	10091- 10210	119	100,3	-15,798	2,372	2,336	4,619	91,695	256,14	0,23		0,158	0,421	19,24	0,123	0,877

Tabla 37. Resultados de evaluación petrofísica del pozo 4

Zonas	Formación	Intervalos	Espesor(ft)	REGISTROS						CÁLCULOS PETROFÍSICOS					
				GR	SP	MINV	MNOR	RESM	RESD	VSH	DPHI	Rw	F	Sw	Sh
Zona 0	...	8700-9370	670	67,114	73,03	4,6805	3,69	37,594	8,193	0,284	0,249	0,505	10,476	0,804	...
Zona 1	CALIZA-A	9370-9477	107	50,96	75,598	12,591	11,39	59,907	22,75	0,199	0,092	0,133	39,685	0,482	...
Zona 2	U SUPERIOR	9477-9537	60	61,906	70,071	5,6763	4,235	24,867	5,294	0,261	0,269	0,561	9,4431	1	...
Zona 3	U INFERIOR	9537-9580	43	61,384	48,897	3,2062	2,661	35,631	61,86	0,277	0,21	0,402	13,177	0,293	0,707
Zona 4	...	9580-9714	134	46,232	79,895	14,689	14,14	77,287	104,3	0,186	0,159	0,279	18,98	0,225	0,775
Zona 5	CALIZA B	9714-9770	56	74,131	66,501	10,748	8,705	20,47	11,95	0,331	0,086	0,124	42,845	0,666	...
Zona 6	T SUPERIOR	9770-9381	61	56,454	53,586	6,886	5,256	14,508	7,45	0,227	0,211	0,405	13,061	0,739	...
Zona 7	T INFERIOR	9381-9879	48	33,837	36,541	3,2801	2,991	77,998	127,1	0,125	0,257	0,528	10,032	0,204	0,796
Zona 8	...	9879-9997	118	104,23	78,508	4,4249	3,603	12,742	6,949	0,548	0,266	0,552	9,597	0,873	...
Zona 9	H SUPERIOR	9997-10032	35	82,486	68,375	9,101	8,453	91,831	29,14	0,397	0,152	0,261	20,252	0,291	...
Zona 10	H INFERIOR	10032-10152	120	42,187	85,373	5,1756	5,479	131,16	122,8	0,176	0,129	0,212	25,018	0,151	0,849

Tabla 38. Resultados de evaluación petrofísica del pozo 5

ZONAS	Formación	Intervalos	Espesor(ft)	REGISTROS						CÁLCULOS PETROFÍSICOS					
				GR	SP	MINV	MNOR	RESM	RESD	VSH	DPHI	Rw	F	Sw	Sh
Zona 0	...	8750-9381	631	93,687	83,552	2,813	2,838	54,687	5,1166	0,296	0,249	0,324	10,498	0,815	...
Zona 1	CALIZA-A	9381-9494	113	70,418	83,237	7,734	7,484	41,16	15,693	0,19	0,124	0,129	26,412	0,465	...
Zona 2	U SUPERIOR	9494-9560	66	84,125	85,153	3,093	2,764	115,27	3,3985	0,257	0,382	0,572	5,9387	1	...
Zona 3	U INFERIOR	9560-9586	26	46,918	22,026	2,217	3,381	1,7489	62,763	0,112	0,206	0,252	13,462	0,233	0,767
Zona 4	...	9586-9724	138	50,358	82,803	7,926	7,853	36,892	62,758	0,119	0,172	0,198	17,152	0,233	...
Zona 5	CALIZA B	9724-9785	61	82,114	43,997	8,388	7,744	155,76	12,973	0,239	0,073	0,063	53,882	0,512	...
Zona 6	T SUPERIOR	9785-9838	53	75,752	28,961	5,204	5,074	1,7823	6,3263	0,21	0,201	0,244	13,934	0,649	...
Zona 7	T INFERIOR	9838-9898	60	49,137	-6,363	2,999	4,135	2,4955	46,93	0,107	0,241	0,311	10,922	0,269	0,731
Zona 8	...	9898-10016	118	134,5	49,927	3,838	3,479	39,959	5,0343	0,567	0,209	0,256	13,267	0,822	...
Zona 9	H SUPERIOR	10016-10085	69	94,302	50,777	6,356	7,049	4,9476	53,623	0,305	0,179	0,209	16,29	0,152	...
Zona 10	H INFERIOR	10085-10130	45	34,019	64,935	1,629	2,592	1,4224	57,356	0,052	0,339	0,489	6,9555	0,339	...