## Universidad Autónoma de Nuevo León

# FACULTAD DE CIENCIAS FÍSICO MATEMÁTICO

#### Maestría en Ciencias de Datos



TÍTULO DE LA TESIS

POR

Juan Jesús Torres Solano

COMO REQUISITO PARCIAL PARA OBTENER EL GRADO DE MAESTRÍA EN CIENCIA DE DATOS

#### Universidad Autónoma de Nuevo León Facultad de Ciencias Físico Matemático

Los miembros del Comité de Tesis recomendamos que la Tesis «Título de la tesis», realizada por el alumno Juan Jesús Torres Solano, con número de matrícula 2173262, sea aceptada para su defensa como requisito parcial para obtener el grado de Maestría en Ciencia de Datos.

El Comité	de Tesis
Dr. José Anastasi	o Hernández Saldaña
A	sesor
Nombre del revisor C	Nombre del revisor D
Revisor	Revisor
Ve	о. Во.
	oxóchitl Ríos Mercado
Subdirector de E	Estudios de Posgrado

Aquí puedes poner tu dedicatoria si es que tienes una.

# ÍNDICE GENERAL

A	grade	ecimientos	III		
1.	. INTRODUCCIÓN . DELIMITACIÓN Y PLANTEAMIENTO DEL PROBLEMA				
2.					
3.	JUS	STIFICACIÓN	3		
4.	FOI	RMULACIÓN DE OBJETIVOS	4		
	4.1.	Generales	4		
	4.2.	Específicos	4		
5.	MA	RCO TEÓRICO	5		
	5.1.	Método Geoléctrico	6		
		5.1.1. Principios Básicos	6		
	5.2.	Sondeo Eléctrico Vertical	6		
		5.2.1. Procesado de Sondeo	6		
		5.2.2. Interpretación Hidrogeológica	6		
	5.3.	Árboles de Decisión	6		
	5.4.	Random Forest	6		

ÍN	ÍNDICE GENERAL V					
		5.4.1.	Número de Árboles	6		
		5.4.2.	Bootstrap	6		
		5.4.3.	Overfitting y Underfitting	6		
		5.4.4.	OOB Score	6		
	5.5.	Valida	ción Cruzada	6		
	5.6.	Métric	eas de Evaluación	6		
6.	6. METODOLOGÍA					
7.	. RESULTADOS Y CONCLUSIONES					
Α.	Apé	éndice	I	9		

# ÍNDICE DE FIGURAS

# ÍNDICE DE TABLAS

#### AGRADECIMIENTOS

Aquí puedes poner tus agradecimientos. (No olvides agradecer a tu comité de tesis, a tus profesores, a la facultad y a CONACyT en caso de que hallas sido beneficiado con una beca).

## Capítulo 1 INTRODUCCIÓN

que es la geofisica:

La exploración geofísica consiste en un conjunto de metodologías que a través de la medición de propiedades petrofisicas del subsuelo es

eL geoelectrica

para que sirve

oportunidades de la cuiencia de datos en el análisis en tiempo real de la respuesta

# Capítulo 2 DELIMITACIÓN Y PLANTEAMIENTO DEL PROBLEMA

Durante un trabajo de prospección geofísica, al realizar adquisición de datos geoelectricos in situ, no es posible conocer el resultado del trabajo hasta una vez realizado procesos de inversión de datos resultado de prospección, por lo que no se tiene certeza de si una lectura presenta inconsistencias o si forma parte de una respuesta esperada para determinado resultado.

El problema es no contar con un modelo que pueda anticipar el resultado de un modelo de inversión geológico, o al menos aproximar un resultado que pueda tener coherencia geológica en interpretación

Como primera aproximación a una solución se considerara un primer supuesto, en el cual se cuenta con respuestas de sondeos eléctricos verticales, más simple pero no menos complejo de interpretar, al contar con un solo segmento de señal, dentro de esta señal podemos encontrar distintas unidades geológicas, profundidad de acuífero, espesores de unidades y profundidad de exploración.

La propuesta de solución es mediante Aprendizaje automático (ML, por sus siglas en inglés), empleando una técnica de aprendizaje que permita identificar patrones que estén asociados a la respuesta de un modelo de inversión, de esta manera obtener un modelo pronóstico de la inversión, sirviendo de guía para incrementar la densidad de muestreo para mejorar el modelo y respuestas.

#### Capítulo 3 JUSTIFICACIÓN

Existen múltiples aplicaciones de ML y DL en el procesamiento, modelado e interpretación geofísica (INSERTAR REFERENCIAS), así como aplicaciones directas por software en las disciplinas mas comerciales (Sísmica de exploración de hidrocarburos PETREL, Shlumberger, INSERTAR REFERENCIA).

Es de esperar que al implementara ciencias de datos mejoraren aspectos básicos, así como especializados, en diversas áreas de las ciencias, mejorando la analítica e interpretación de resultados, o bien, modelando predicciones.

En este sentido es oportuna la implementación de modelos de regresión y clasificación durante la adquisición de datos de un Sondeo Eléctrico Vertical (SEV, VES por sus siglas en ingles), el proceso de adquisición in situ consta de un intervalo de muestreo predefinido, el cual esta acotado de acuerdo al objetivo de exploración, siendo el principal factor establecer un intervalo de muestreo que permita identificar las unidades geológicas de interés, es decir mantener un intervalo de muestreo menor a la frecuencia de ocurrencia de las unidades, por lo que el éxito de la exploración dependerá en su totalidad de la planeación previa de la adquisición.

De manera que un modelo que permita clasificar las lecturas y generar una regresión para proponer muestreos adicionales *in situ* tendría la ventaja de optimizar y mejorar la calidad de la adquisición y muestreo de las frecuencias deseadas.

#### Capítulo 4 FORMULACIÓN DE OBJETIVOS

#### 4.1 Generales

- Generar un modelo predictivo que permita clasificar datos geoeléctricos durante la adquisición de SEV's in situ.
- Establecer un modelo de entrenamiento de regresión mediante Random Forest que genere intervalos de muestreo en SEV's al identificar contrastes en los datos.

#### 4.2 Específicos

- Establecer modelos de geoeléctricos con unidades geológicas de terrenos ya caracterizados mente SEV's.
- Generar un modelo de clasificación y regresión mediante Random Forest mediante modelos geoeléctricos.

# Capítulo 5

## MARCO TEÓRICO

- 5.1 MÉTODO GEOLÉCTRICO
- 5.1.1Principios Básicos
  - 5.2 Sondeo Eléctrico Vertical
- 5.2.1Procesado de Sondeo
- 5.2.2Interpretación Hidrogeológica
  - 5.3 ÁRBOLES DE DECISIÓN
    - 5.4 Random Forest
- 5.4.1Número de Árboles
- 5.4.1.1Max Tree Depth
- 5.4.2Bootstrap
- 5.4.3 Overfitting y Underfitting
- 5.4.4OOB Score
- 5.5 VALIDACIÓN CRUZADA
- 5.6 MÉTRICAS DE EVALUACIÓN

## Capítulo 6 METODOLOGÍA

## Capítulo 7 RESULTADOS Y CONCLUSIONES

## Apéndice A Apéndice I

## Bibliografía