

APLICACIÓN DEL ANÁLISIS MULTIVARIADO Y REDES NEURONALES  
ARTIFICIALES PARA ESTIMAR EL MÓDULO DE YOUNG EN SHALES

JUAN PABLO ESPINOSA CARDENAS

UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍAS FISICOMECHANICAS  
ESCUELA DE INGENIERÍA CIVIL  
BUCARAMANGA

2014

APLICACIÓN DEL ANÁLISIS MULTIVARIADO Y REDES NEURONALES  
ARTIFICIALES PARA ESTIMAR EL MÓDULO DE YOUNG EN SHALES

Ing. JUAN PABLO ESPINOSA CARDENAS

Trabajo de grado presentado como requisito para optar al título de  
MAGISTER EN GEOTECNIA

Director

OSCAR JAVIER BEGAMBRE CARRILLO. IC, MSc, Ph.D

Co-Director

M. Sc. JENNY MABEL CARVAJAL JIMENEZ

UNIVERSIDAD INDUSTRIAL DE SANTANDER  
FACULTAD DE INGENIERÍAS FISICOMECHANICAS

ESCUELA DE INGENIERÍA CIVIL

BUCARAMANGA

2014

## DEDICATORIA

Al Amor de mi vida: Cindy Guerrero. A las personas llenas de valores y amor infinito: mis papas Carmen Elisa y Laureano. A mis hermanos y amigos Leidy Andrea y Julian, y al futuro de la familia: mis sobrinos Santiago y Thomas.

## **AGRADECIMIENTOS**

A Dios, por darme la vida, la salud y el carácter para permanecer firme en mis sueños.

A Cindy Mildred por su amor y apoyo. Por darle ese toque especial a mi vida y llenarla de alegría.

A mis padres Carmen Elisa y Laureano por su apoyo, motivación y ejemplo para formarme como persona íntegra, con valores que son la base de mis triunfos.

A mis hermanos Julian y Leidy Andrea por su confianza, por expresarme siempre su admiración, que hace que me mantenga motivado.

A mi director Oscar Begambre, por su apoyo y sabiduría. Por la confianza y calidad depositada en esta tesis.

A mi codirectora, Jenny Carvajal, por enfocarme en este estudio, por sus brillantes ideas y su colaboración.

Al equipo de Geomecánica y Laboratorio de mecánica de rocas-ICP por permitirme el uso de información técnica que hizo posible llevar a cabo este trabajo.

## TABLA DE CONTENIDO

<b>INTRODUCCION.....</b>	<b>15</b>
<b>1 ANÁLISIS MULTIVARIADO Y SU APLICACION EN LA GEOTECNIA Y GEOMECAÁNICA.....</b>	<b>17</b>
1.1 Técnicas de Análisis Multivariado.....	17
1.1.1 Análisis Clúster:.....	17
1.1.2 Análisis discriminante:.....	17
1.1.3 Componentes Principales .....	17
1.1.4 Regresión Múltiple.....	18
1.2 Aplicaciones de la Regresión Múltiple para Determinar Propiedades Mecánicas de las Rocas: Revisión de la Literatura. ....	22
<b>2 REDES NEURONALES APLICADAS A SHALES.....</b>	<b>27</b>
2.1 Reconocimiento estadístico de patrones (REP).....	27
2.2 Redes neuronales artificiales.....	28
2.3 Predicción de Propiedades Mecánicas de la Roca Utilizando Redes Neuronales: Revisión de la Literatura. ....	31
2.4 MODULO DE YOUNG VS PROPIEDADES PETROFISICAS .....	33
<b>3 VALIDACIÓN DEL ANÁLISIS MULTIVARIADO Y LAS REDES NEURONALES. ....</b>	<b>35</b>
3.1 Análisis Multivariado .....	35
3.1.1 Correlaciones para la Resistencia Compresiva Uniaxial (UCS) y el Modulo de Young (E) determinadas en Artículo (DEGHAN, SATTARI, CHEHREH; 2010).....	36
3.1.2 Correlación para el cálculo del UCS:.....	38
3.1.3 Correlación para Cálculo de E.....	38
3.1.4 Análisis de varianza (ANOVA) para comparación de resultados obtenidos en este estudio con el trabajo de DEGHAN, SATTARI, CHEHREH (2010).....	39
3.1.5 Tratamiento Estadístico para mejorar las correlaciones en el Cálculo del UCS y Young. ....	43
3.2 Redes Neuronales en el cálculo del UCS y Modulo de Young a partir de otros ensayos (porosidad, carga puntual, velocidad de Onda compresional y Numero de rebotes del martillo Schmidt .....	48
3.2.1 UCS utilizando redes neuronales.....	48
3.2.2 Módulo de Young Utilizando Redes Neuronales.....	49
<b>4 APLICACIÓN DEL ANÁLISIS MULTIVARIADO A SHALES PARA EL CÁLCULO DEL MÓDULO DE YOUNG (E).....</b>	<b>52</b>
4.1 Regresión Múltiple Pozo La luna.....	58

4.2	Test de Grubbs formación la luna miembro Galembo, Orientación horizontal. ....	61
4.3	Identificación de variables estadísticamente significativas:.....	64
<b>5</b>	<b>APLICACIÓN DE LAS REDES NEURONALES A SHALES PARA EL CÁLCULO DEL MÓDULO DE YOUNG (E).....</b>	<b>67</b>
<b>6</b>	<b>ANÁLISIS DE RESULTADOS .....</b>	<b>71</b>
6.1	RESULTADOS ANALISIS MULTIVARIADO .....	71
6.1.1	Metodología para determinar Parámetros geomecánicos a Partir de Registros petrofísicos y pruebas de laboratorio.....	71
6.1.2	Ventajas del Análisis Multivariado.....	73
6.2	RESULTADOS CON LAS REDES NEURONALES.....	74
<b>7</b>	<b>CONCLUSIONES .....</b>	<b>75</b>
<b>8</b>	<b>RECOMENDACIONES.....</b>	<b>77</b>
	<b>BIBLIOGRAFÍA.....</b>	<b>78</b>
	<b>ANEXOS.....</b>	<b>80</b>

## LISTA DE FIGURAS

Figura 1-1 Representación gráfica del método de los mínimos cuadrados .....	19
Figura 2-1. Autómatas Celulares.....	28
Figura 2-2. Neurona artificial (McCulloch-Pitts) .....	29
Figura 2-3 Función Sigmoide .....	29
Figura 2-4 Taxonomía de RNA .....	30
Figura 2-5 Arquitectura de GRNN.....	32
Figura 2-6. Módulo de Young Vs Contenido de arcilla .....	34
Figura 2-7. Módulo de Young Vs Porosidad .....	34
Figura 3-1 Curva de frecuencias módulo de Young .....	44
Figura 3-2. Curva de frecuencias Resistencia Compresiva Uniaxial (UCS) .....	44
Figura 4-1.Roca shale .....	52
Figura 4-2. Gráfico de frecuencia del Módulo de Elasticidad-formación la luna, miembro Galembo .....	61
Figura 5-1. Interface de usuario principal ORION .....	67
Figura 5-2. Interface de carga de Datos .....	68

## LISTA DE TABLAS

Tabla 3-1. Resultados experimentales de las pruebas para la determinación de E y UCS.....	35
Tabla 3-2 Variables independientes modelo de Regresión multivariado .....	37
Tabla 3-3 Estadística de la Regresión de UCS para cada Variable. ....	38
Tabla 3-4. Estadística de la Regresión del Módulo de Young para cada Variable.....	38
Tabla 3-5. Datos Observados y resultados de las ecuaciones. ....	39
Tabla 3-6. ANOVA módulo de Young de datos observados y regresión artículo de DEHGHAN, SATTARI, CHEHREH (2010).....	41
Tabla 3-7. ANOVA UCS de datos observados y regresión artículo de DEHGHAN, SATTARI, CHEHREH (2010) .....	41
Tabla 3-8. ANOVA módulo de Young de datos observados y regresión estudio .....	41
Tabla 3-9. ANOVA UCS de datos observados y regresión estudio .....	42
Tabla 3-10. ANOVA módulo de Young de Regresión Artículo (DEHGHAN, SATTARI, CHEHREH ;2010) y regresión estudio .....	42
Tabla 3-11. ANOVA UCS de Regresión Artículo (DEHGHAN, SATTARI, CHEHREH ;2010) y regresión estudio .....	43
Tabla 3-12 Resultado de Cálculo de T.....	45
Tabla 3-13. Estadística de la Regresión para UCS aplicando Test de Grubbs .....	46
Tabla 3-14. Análisis de varianza UCS aplicando Test de Grubbs .....	46
Tabla 3-15. Estadística de la Regresión para( E) aplicando Test de Grubbs .....	47
Tabla 3-16. Análisis de varianza (E) aplicando Test de Grubbs .....	47
Tabla 3-17. Coeficiente de determinación para diferentes Configuración de redes neuronales .....	48
Tabla 3-18. Datos utilizados en red neuronal mejor coeficiente de Determinación .....	49
Tabla 3-19. Coeficiente de determinación y Configuración de redes .....	50
Tabla 3-20. Datos entrenamiento y red con mejor ajuste en el módulo de Young .....	50
Tabla 4-1 Densidad de datos módulo Elasticidad .....	53
Tabla 4-2. Datos de Registros formación La Luna .....	54
Tabla 4-3. Coeficientes de Determinación Pozo la Luna 1 .....	56
Tabla 4-4. Coeficientes de Determinación Formación la Luna horizontal.....	57
Tabla 4-5. Coeficientes de Determinación miembro Galembo, orientación horizontal. ....	57
Tabla 4-6. Coeficientes de variación miembro Galembo, orientación Vertical .....	57
Tabla 4-7. Estadística de la Regresión pozo la luna, orientación vertical.....	58
Tabla 4-8. Análisis de varianza pozo la luna, orientación vertical.....	58
Tabla 4-9. Estadística de la Regresión Formación la luna, orientación Horizontal .....	59
Tabla 4-10. Análisis de varianza Formación la luna, orientación Horizontal .....	59
Tabla 4-11. Estadística de la Regresión Formación la luna, Miembro Galembo, orientación Horizontal .....	60

Tabla 4-12. Análisis de Varianza Formación la luna, Miembro Galembo, orientación Horizontal.....	60
Tabla 4-13. Estadística de la Regresión Formación la luna, Miembro Galembo, orientación Vertical. ....	60
Tabla 4-14. Análisis de varianza Formación la luna, Miembro Galembo, orientación Vertical.....	61
Tabla 4-15. T calculados data Pozo la Luna.....	62
Tabla 4-16. Estadística de la Regresión con 95% de probabilidad Test de Grubbs ....	63
Tabla 4-17. Análisis de variables con 95% de probabilidad Test de Grubbs.....	63
Tabla 4-18. Estadística de la Regresión con 95% de probabilidad variables independientes y 90 % variable dependiente en Test de Grubbs .....	64
Tabla 4-19. Análisis de variables con 95% de probabilidad variables independientes y 90 % variable dependiente en Test de Grubbs.....	64
Tabla 4-20. Estadística de la Regresión eliminando dos variables no significativas ....	65
Tabla 4-21. Análisis eliminando dos variables no significativas.....	65
Tabla 4-22. Estadística de la Regresión eliminando tres variables no significativas..	65
Tabla 4-23. Análisis eliminando tres variables no significativas.....	65
Tabla 5-1. Coeficientes de Determinación diferentes redes neuronales .....	69
Tabla 5-2. Datos de entrenamiento y prueba con coeficiente de determinación más alto. ....	69
Tabla 6-1. Comparación coeficientes de Variación entre regresión Múltiple y Redes Neuronales.....	74
Tabla 7-1. Cuadro Comparativo de requerimientos entre las técnicas de análisis Multivariado y Redes Neuronales. ....	76

## LISTA DE ANEXOS

<b>Anexo A:</b> Graficas Regresión Simple Pozo La luna-1, Vertical. Formación la Luna, Simití y tablazo.....	80
<b>Anexo B:</b> Graficas Regresión simple Formación La Luna del pozo la luna-1, Horizontal.....	82
<b>Anexo C:</b> Graficas Regresión Simple Formación la luna miembro Galembo Horizontal.....	85
<b>Anexo D:</b> Graficas Regresión Simple Formación La Luna, miembro Galembo, Orientación Vertical.....	88

## RESUMEN

TÍTULO: “APLICACIÓN DEL ANÁLISIS MULTIVARIADO Y REDES NEURONALES ARTIFICIALES PARA ESTIMAR EL MÓDULO DE YOUNG EN SHALES”.\*

AUTOR: JUAN PABLO ESPINOSA CARDENAS \*\*

**Palabras claves:** modelos predictivos, shales, Modulo de Young, análisis multivariado, redes neuronales.

Los modelos geomecánicos se construyen a partir de registros, cuya medición se realiza en campo de forma indirecta. Para tener confianza en estos modelos se hace necesario calibrarlos a partir de datos obtenidos a través de ensayos de laboratorio.

En muchas ocasiones contar con buenos datos de laboratorio, distribuidos a diferentes profundidades, no es muy común debido a la dificultad en la preparación de las muestras y en la medición de las propiedades durante la ejecución y análisis de los ensayos. Cuando las formaciones a evaluar son tipo shale los problemas aumentan, debido a algunas propiedades particulares de estas rocas como son su fisilidad, la presencia de planos de debilidad y fracturas.

En aras de resolver estos inconvenientes se recurre a los modelos predictivos con el objeto de obtener datos suficientes para poblar los modelos geomecnicos y calibrarlos. Se busca así aumentar la confiabilidad para la determinación del Módulo de Young en formaciones de shale.

La metodología utilizada en este trabajo consiste en una serie de análisis a datos provenientes de laboratorio y de registro de pozo. Inicialmente se establece, a través de un análisis multivariado, una relación entre variables petrofísicas (Gamma Ray, porosidad, densidad, resistividad, tiempos de transito de ondas P y S,) y el módulo de Young (obtenido de laboratorio) en shales. Posteriormente, se desarrolla una técnica a partir de redes neuronales artificiales utilizando los mismos datos de etapa anterior. Finalmente se evalúa y compara el desempeño de la red neuronal con el análisis multivariado en términos de su correlación y confiabilidad en la determinación del módulo de elasticidad.

Los Resultados mostraron que utilizando las Redes Neuronales se obtienen mayores coeficientes de determinación  $R^2$  pero a un costo más alto que si se aplica el análisis multivariado. Por tanto se ahorra tiempo y dinero con esta última técnica.

---

\* Trabajo de Grado

\*\* Facultad de Ingenierías Físico – Mecánicas. Escuela de Ingeniería civil  
Director: Ph.D. Oscar Javier Begambre Carrillo, Co-director: M. Sc. Jenny Mabel Carvajal Jimenez,

## ABSTRACT

TITLE: "APPLICATION OF MULTIVARIATE ANALYSIS AND ARTIFICIAL NEURAL NETWORKS TO PREDICT YOUNG'S MODULUS OF SHALE".<sup>1</sup>

AUTHOR: JUAN PABLO ESPINOSA CARDENAS \*\*

**keywords:** predictive models, shales, young's modulus, multivariate analysis , artificial neural networks.

Geomechanical models are built from logs, as measured on field indirectly. To have reliability in these models it is necessary to calibrate using data obtained through laboratory tests.

Often have good laboratory data distributed at different depths, is not very common due to the difficulty in sample preparation and measurement of the properties during the execution and analysis of tests. When evaluating formations are shale type, the problems increase, due to some particular properties of these rocks as are his fisilidad, the presence of fractures and planes of weakness

In order to solve these drawbacks uses predictive models in order to obtain sufficient data to populate and calibrate the geomechanical models. It seeks to increase the reliability to determine the Young's Modulus in shale formations.

The methodology used in this work consists of a series of analyzes to data from laboratory and well log. Initially established through multivariate analysis, a relationship between petrophysical variables (Gamma Ray, porosity, density, resistivity, transit times of P and S waves) and Young's modulus (obtained from laboratory) in shales. Subsequently, a technique developed from artificial neural networks using the same data of previous stage. Finally it evaluates and compares the performance of the neural network with multivariate analysis in terms of their correlation and reliability in determining the modulus of elasticity.

Results show that neural networks using higher coefficients of determination  $R^2$  are obtained, but at a higher cost than if the multivariate analysis is applied. Therefore, saves time and money with this technique last, and predictions show good correlation.

---

<sup>1</sup> Graduate Project-Thesis

\*\* Faculty of physic – Mechanical engineering, civil engineering schools  
Director: Ph.D. Oscar Javier Begambre Carrillo, Co-director: M. Sc. Jenny Mabel Carvajal Jimenez,

## INTRODUCCIÓN

El módulo de Young es uno de los parámetros más importantes, requeridos y determinados en estudios de mecánica de rocas en diferentes campos para proyectos de minería, ingeniería civil y de ingeniería de petróleos. En operaciones de perforación de pozos tiene sus aplicaciones principalmente para la estimación de ventanas operacionales de estabilidad, colapso de tuberías de revestimiento, localización del pozo y diseño de trayectorias.

Específicamente en el análisis de estabilidad de un pozo se requiere caracterizar mecánicamente la formación, determinando entre otros los parámetros elásticos, el módulo de Young y relación de poisson, la resistencia a la compresión, ángulo de fricción interna y cohesión para construir un modelo de roca. Este modelo puede ser construido empleando algún software comercial a partir de registros, pero al ser éstos una medición indirecta (realizada, por ejemplo, con velocidades de onda o tiempos de tránsito), el modelo geomecánico podría no ajustarse al comportamiento real y requiere ser calibrado.

Por tanto son de gran importancia los valores de los parámetros mecánicos determinados a partir de mediciones directas en laboratorio porque permiten calibrar las curvas generadas en el modelo geomecánico. Cuando se dispone de varios datos a diferentes profundidades se constituye un “modelo de roca” que es fundamental para la determinación de la ventana operacional.

El procedimiento de medida de los parámetros elásticos han sido estandarizados por la ASTM<sup>2</sup> y el ISRM<sup>3</sup>. Muestras de alta calidad son necesarias para llevar a cabo los ensayos en el laboratorio. En el proceso de preparación una gran cantidad de muestra se daña, particularmente los Shale que son fisiles debido a la presencia planos de debilidad y fracturas. Además, llevar a cabo la ejecución de estos ensayos con éxito no es fácil, requiere especial cuidado y destreza, principalmente en el cálculo de la deformación necesaria para la obtención del módulo de Young y la relación de Poisson. En efecto, la tasa de alcance de resultados exitosos es baja y obtener módulos de elasticidad en el laboratorio se convierte en un gasto de tiempo y dinero, que en muchos casos las empresas no quieren costear, pero realmente es necesario poblar los modelos geomecánicos con datos de laboratorio para poder calibrarlos.

Para superar estas dificultades durante la preparación de las muestras y la ejecución de los ensayos, y su alto costo se puede emplear los modelos

---

<sup>2</sup> American Society for Testing and Materials: Standard Test Method For Compressive Strength And Elastic Moduli Of Intact Rock Core Specimens Under Varying States Of Stress And Temperatures. ASTM D7012, united States, 2010.

<sup>3</sup> International Society for rock Mechanics: Determining the Uniaxial Compressive Strength and Deformability of Rock Materials. EUR, 1979.

predictivos, ya que su uso reduce la cantidad de muestras a ensayar, bajando el costos y el tiempo de respuesta de resultados y además se obtienen cualquier cantidad de datos que permiten poblar y calibrar los modelos.

Tradicionalmente se han utilizado técnicas estadísticas como la regresión lineal simple y la regresión lineal múltiple para obtener modelos predictivos en ingeniería de rocas. En los últimos años otros métodos como las redes neuronales artificiales y la lógica difusa han incursionado como herramientas de predicción. (Ver trabajo de DEGHAN, SATTARI, CHEHREH, en 2010)

Este trabajo está enfocado al estudio de modelos predictivos con el objeto de reducir la incertidumbre en el modelamiento de estabilidad de pozos específicamente en formaciones de Shale para obtener el Módulo de Elasticidad. La metodología utilizada consiste en una serie de análisis de datos provenientes de laboratorio y de registro. Inicialmente, se establece, a través de un análisis multivariado, una relación entre variables petrofísicas (porosidad, densidad, resistividad, Gamma Ray, Tiempo de transito ondas P y S) provenientes de registros y de ensayos de laboratorio y el módulo de Young en shales. Posteriormente se desarrolla una técnica a partir de redes neuronales artificiales para estimar el módulo de Young en shales utilizando datos disponibles (registros) y Modulo de Young (Laboratorio). finalmente se evalúa y compara el desempeño de la red neuronal con el análisis multivariado en términos de su correlación y confiabilidad en la determinación del módulo de elasticidad.

# 1 ANÁLISIS MULTIVARIADO Y SU APLICACIÓN EN LA GEOTECNIA Y GEOMECÁNICA.

Las técnicas multivariadas son poderosas herramientas para ayudar en la evaluación de la información. En el campo Geológico se puede reconocer la aplicación de técnicas multivariadas desde 1830. El reconocimiento de que existen diferencias entre grupos de objetos constituye la base de tales análisis estadísticos. Sin embargo hay una gran diferencia entre el reconocimiento de diferencias y un intento por formular declaraciones probabilísticas en cuanto a la importancia de la diferencia observada.

## 1.1 Técnicas de Análisis Multivariado

Entre las técnicas más utilizadas se encuentra:

1. Regresión Múltiple
2. Análisis Clúster
3. Análisis Discriminante
4. Componentes principales

### 1.1.1 Análisis Clúster:

Es un amplio término genérico para una multitud de procedimientos los cuales son diseñados para separar un conjunto de objetos dentro de grupos relacionados o familias.<sup>4</sup>

### 1.1.2 Análisis discriminante:

Cuando se sabe que los datos vienen de más de una fuente y se desea ser capaz de reconocer que fuente da lugar a los datos entonces un enfoque estadístico llamado Análisis discriminante es apropiado. Un criterio de un conjunto de datos es usado para establecer una o más ecuaciones las cuales son derivadas de tal forma que se encuentre un plano o hiperplano que separe mejor las diferentes fuentes. El número de planos de separación es el mínimo del número de variables y el mínimo número de fuentes (grupos).<sup>5</sup>

### 1.1.3 Componentes Principales

Es un método ampliamente utilizado en estadística par el análisis de datos, se ocupa de explicar la estructura de varianza-Covarianza de un conjunto de variables por medio de una combinación lineal de dichas variables. Sus objetivos principales son Reducción e interpretación de datos<sup>6</sup>. Por tanto en un banco de datos de

---

<sup>4</sup> Ibid., p.5.

<sup>5</sup> Ibid., p.6.

<sup>6</sup> SOLIS, G y MARROQUIN, E. Aplicación de Técnica de Análisis Estadístico Multivariado en un proceso de control Dimensional, IEEE LATIN AMERICA TRANSACTIONS, Vol. 5, mayo, 2007.

muchas variables la finalidad será reducirlas a un menor número perdiendo la menor cantidad de información posible.

Los componentes principales nuevos serán una combinación lineal de las variables originales, y a su vez estas serán independientes entre sí.

#### 1.1.4 Regresión Múltiple

Cuando se aplica la regresión múltiple, en realidad se construye un modelo que explica la variabilidad de la variable dependiente. Para eso se incorporan las influencias simultáneas e individuales de varias variables independientes. En regresión múltiple se busca obtener una ecuación de la forma:

$$Y = a_0 + \sum_{i=1}^N a_i X_i \quad \text{Ecuación 1-1}$$

En donde

Y: es el valor predicho de la variable de interés.

a: son un set de coeficientes a ser encontrados

X: son las variables observadas bajo las cuales la Y es aparentemente dependiente

Para cada valor observado de Y, la variable de interés, existen observaciones en las x's. La solución para las a's es encontrada con la resolución de un set de ecuaciones como:

$$Q = \sum (Y_i - \hat{Y}_i)^2 \quad \text{Ecuación 1-2}$$

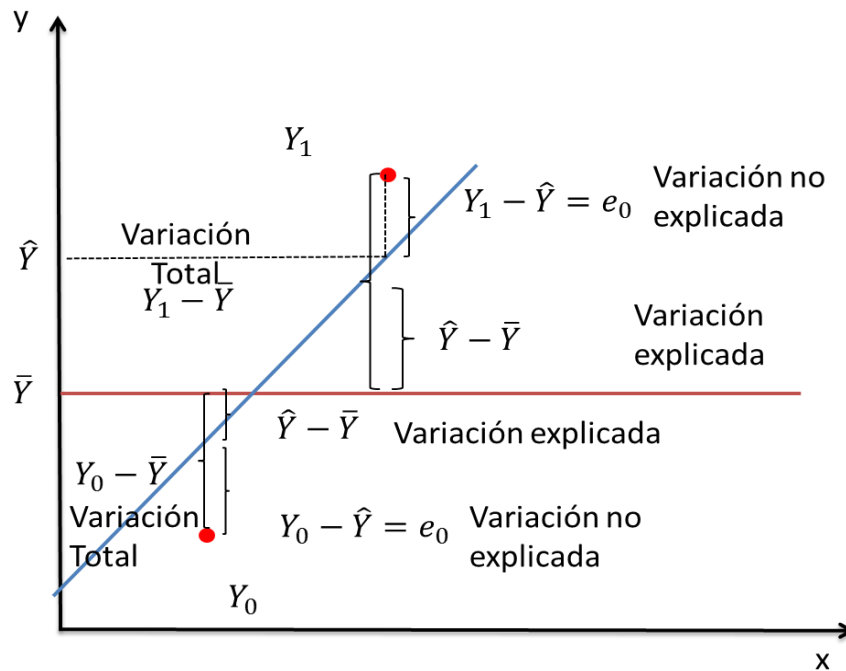
Donde Q es un mínimo en un caso más general se puede representar como  $\hat{Y} = f(x)$  donde  $f(x)$  necesariamente no es lineal<sup>7</sup>

##### 1.1.4.1 Método de los mínimos cuadrados

La regresión se basa en el método de los mínimos cuadrados. A continuación se muestra una representación gráfica del mismo.

---

<sup>7</sup> Ibid., p.4.



**Figura 1-1** Representación gráfica del método de los mínimos cuadrados  
**Modificado de:** Estadística para administración y economía, 2008.

Como se observa en la gráfica, cada dato ( $y$ ) es igual a este mismo punto reflejado en la línea azul ( $\hat{Y}$ ), más un error. Matemáticamente se expresa:

$$Y_i = \hat{Y}_i + e_i \quad \text{Ecuación 1-3}$$

Si se resta la media muestral a ambos miembros se obtiene:

$$\begin{aligned} Y_i - \bar{Y} &= \hat{Y}_i + e_i - \bar{Y} \\ Y_i - \bar{Y} &= (\hat{Y}_i - \bar{Y}) + (Y_i - \hat{Y}_i) \end{aligned}$$

Que se puede interpretar de la manera siguiente:

*la desviación observada con respecto a la media muestral*  
 = *a la desviación predicha con respecto a la media muestral*  
 + *Residuo*

$$\begin{aligned} \sum_{i=1}^n (Y_i - \bar{Y}) &= \sum_{i=1}^n (\hat{Y}_i - \bar{Y}) + (Y_i - \hat{Y}_i) \\ &= \sum_{i=1}^n (\hat{Y}_i - \bar{Y}) + \sum_{i=1}^n e_i \end{aligned}$$

De esta forma se llega a la descomposición de los mínimos cuadrados que se conoce típicamente:

$$STC = STR + SCE$$

Ecuación 1-4

Suma total de los cuadrados=suma de los cuadrados de la regresión +suma de los cuadrados de los errores.

#### 1.1.4.2 Descomposición de la suma de los cuadrados y el coeficiente de determinación.<sup>8</sup>

El modelo de regresión múltiple ajustado mediante mínimos cuadrados se puede representar:

$$Y_i = b_0 + b_1 x_{1i} + b_2 x_{2i} + \dots + b_k x_{ki} + e_i = \hat{Y}_i + e_i$$

En el modelo las  $b_i$  son las estimaciones por mínimos cuadrados de los coeficientes del modelo de regresión poblacional y las  $e$  son los residuos del modelo de regresión estimado. La variabilidad del modelo se puede dividir en tres diferentes componentes que son  $STC$ ,  $SCR$  y  $SCE$  que se mencionaron anteriormente y corresponden a la ecuación 1-4

Donde estos componentes se definen de la siguiente manera:

$$STC = \sum_{i=1}^n (Y_i - \bar{Y})^2 = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2 + \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

Ecuación 1-5

Así mismo la suma de los cuadrados de los errores es

$$SCE = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2 = \sum_{i=1}^n e_i^2$$

Ecuación 1-6

La suma de los cuadrados de la regresión:

$$SCR = \sum_{i=1}^n (\hat{Y}_i - \bar{Y})^2$$

Ecuación 1-5

Otra forma para denominar esta descomposición es:

Variabilidad Muestral Total = Variabilidad Explicada + Variabilidad no Explicada.

En la gráfica anterior, figura 2-1 se muestra a que corresponde cada uno de estos componentes.

#### 1.1.4.3 Coeficiente de Determinación, $R^2$

El coeficiente de determinación,  $R^2$ , de la regresión es la proporción de variabilidad muestral total explicada por la regresión<sup>9</sup>. Se utiliza el  $R^2$ , habitualmente como

<sup>8</sup> NEWBOLD, Paul. Estadística para Administración y Economía. 6 ed. Madrid: Pearson Educación, 2008. P. 494

estadístico descriptivo para describir la fuerza de la relación lineal entre las variables independientes X y la variable dependiente Y.

$$R^2 = \frac{SCR}{STC} = 1 - \frac{SCE}{STC} \quad \text{Ecuación 1-6}$$

Se deduce que,

$$0 \leq R^2 \leq 1$$

#### 1.1.4.4 Estimación de la Varianza de errores

La estimación de la varianza de los errores del modelo poblacional se calcula a partir de la suma de los cuadrados los errores.

$$S_e^2 = \sum_{i=1}^N \frac{e_i^2}{N-1-k} \quad \text{Ecuación 1-9}$$

$$S_e^2 = \sum_{i=1}^N \frac{(Y - \hat{Y})^2}{N - 1 - k}$$

En el cual k es el número de variables independientes en el modelo de regresión.

#### 1.1.4.5 Error típico de estimación

El error típico se calcula como la raíz cuadrada de la varianza de los errores.

$$S_e = \sqrt{\frac{\sum_{i=1}^n e_i^2}{n-K-1}} = \sqrt{\frac{SCE}{n-k-1}} \quad \text{Ecuación 1-7}$$

#### 1.1.4.6 Cuadrado Medio de la Regresión

El cuadrado medio de la regresión se expresa matemáticamente como:

$$CMR = \frac{SCR}{K} \quad \text{Ecuación 1-8}$$

El *CMR* se utiliza como medida de la variabilidad explicada ajustada con el que se tienen en cuenta el número de variables independientes.

#### 1.1.4.7 Coeficiente de Determinación Ajustado, $\bar{R}^2$ ,<sup>10</sup>

En la Ecuación 1-8, vemos que el valor de  $R^2$  puede ser alto cuando *SCE* es pequeño, esto debido a que los puntos observados están cerca de los puntos predichos; o  $R^2$  es alto también debido a que el *STC* es grande.

<sup>9</sup> Ibid., p.506.

<sup>10</sup> Ibid., p.50p.

La utilización de  $R^2$ , como una medida global de la calidad de una ecuación ajustada puede plantear un problema. Cuando se suman variables independientes a un modelo de regresión múltiple, en la mayoría de los casos, la suma explicada de los cuadrados, SCR, aumenta así la variable adicional no le da peso a la predicción. Por tanto es común encontrar que  $R^2$  ha aumentado notoriamente después que se han añadido una o más variables de poca importancia predictiva al modelo. Pero en este caso el aumento del valor de  $R^2$ , sería un resultado engañoso. Para superar este inconveniente el valor del **Coefficiente de Determinación Ajustado**  $\bar{R}^2$  puede calcularse con la ecuación siguiente:

$$\bar{R}^2 = \frac{\frac{SCE}{(n - K - 1)}}{\frac{STC}{n - 1}}$$

Se utiliza este parámetro con el objeto de tener en cuenta el hecho de que las variables independientes irrelevantes provocan una pequeña reducción de la suma de los cuadrados, por tanto el  $\bar{R}^2$  ajustado permite comparar mejor los modelos de regresión múltiple, debido a que estos tienen diferentes números de variables independientes.

#### 1.1.4.8 Coeficiente de correlación Múltiple

El coeficiente de correlación múltiple es un indicador de la correlación que existe entre el valor predicho y el valor observado.

$$R = r(\hat{Y}, Y) = \sqrt{R^2}$$

El coeficiente de correlación múltiple es igual a la raíz cuadrada del coeficiente múltiple de determinación. Se utiliza R como otra medida de la fuerza de la relación de la variable dependiente y las variables independientes. Por esto se compara a la correlación entre y y x de la regresión simple.

## 1.2 Aplicaciones de la Regresión Múltiple para Determinar Propiedades Mecánicas de las Rocas: Revisión de la Literatura.

Algunos de los trabajos enfocados en la industria del petróleo para determinar propiedades mecánicas de la roca usando la técnica de análisis multivariado se presentan a continuación:

## **Predicción de la Anisotropía de La resistencia en Compresión Uniaxial (UCS) de Algunas Areniscas Laminadas Usando un Análisis de Regresión Multivariado.<sup>11</sup>**

El trabajo trata Sobre la estimación de la anisotropía del UCS teniendo en cuenta como variables independientes: ángulos de orientación ( $\beta$ ), peso unitario ( $\gamma$ ), y porosidad ( $\eta$ ) usando la técnica de análisis de regresión multivariada, el método requiere al menos 3 valores de UCS de  $\beta=0^\circ$ ,  $30^\circ$  y  $90^\circ$  y la ecuación es válida solamente para un único set de datos o rango de resistencia representativa.

Recientemente en un trabajo anterior, se aplicó las técnicas redes neuronales y regresión lineal multivariada para la predicción del UCS, así como la resistencia tensil y la resistencia de carga puntual de una roca esquistosa a partir de la composición mineral y propiedad textural. El producto obtenido fue la predicción de resultados confiables con el uso de las redes neuronales, pero se obtuvieron valores de correlaciones y resultados inconsistentes desde el modelo estadístico. El problema fue que los parámetros de calidad estadísticos no fueron suficientes. El trabajo de Garagon y Can presenta el comportamiento anisotrópico de la resistencia y su relación con algunas de las propiedades básicas petrofísicas de una arenisca seleccionada. Debido a la existencia de las correlaciones no lineales entre UCS y Porosidad, peso unitario y ángulo de orientación; el aporte de cada variable independiente fue incluido en el análisis de la regresión y de esta forma un modelo de regresión confiable fue desarrollado.

En la primera etapa de la evaluación estadística una serie de análisis de regresión bivariada fueron desarrollados.

**Modelo de Regresión:** Todos los modelos no lineales junto con los modelos lineales fueron probados para ajustar los datos con el objeto de establecer un modelo de regresión bivariado para el UCS usando variables independientes como la porosidad, peso unitario y ángulo de orientación.

El coeficiente de determinación  $R^2$ , el coeficiente ajustado de determinación  $Ra^2$ , y el error estándar de estimación SEE fueron usados como medida numérica de la bondad del ajuste del modelo de regresión.

## **PREDICCIÓN DE LA RESISTENCIA COMPRESIVA UNIAxIAL Y MODULOS DE ELASTICIDAD PARA MUESTRAS TRAVERTINO USANDO REGRESIÓN Y REDES NEURONALES ARTIFICIALES<sup>12</sup>**

Este trabajo tiene dos enfoques que son regresión multivariada y Redes neuronales. **En cuanto a la regresión:** para establecer modelos predictivos entre

---

<sup>11</sup> GARAGON, Mert y CAN Tolga. Predicting the Strength anisotropy in uniaxial compression of some laminated Sandstone using multivariate regression analysis. *Materials and structures* (2010) 43: 509-5017.

<sup>12</sup> DEGHAN, S; SATTARI, Gh; CHEHREH, Chelganis y ALIABADI Ma. Prediction of uniaxial compressive strength and modulus of elasticity for travertine sample using regression and artificial neural networks. *Mining Science and technology* 20 (2010) 0041-0046

las propiedades mecánicas de la roca, el análisis por regresión es un método tradicional, algunos modelos simples usan el valor del martillo Smith, la velocidad de Onda P, la carga puntual y otros que han sido aplicados para estimar UCS y Young.

Sin embargo quien ha establecido modelos simples únicamente teniendo en cuenta el  $R^2$ , sabrá que no es suficiente, porque el  $R^2$  basado en modelos simples no explica la variación total introducida por la variable independiente. Por lo tanto el uso de modelos por regresión múltiple será más exacto en la predicción de propiedades elásticas de la roca. El método puede ser utilizado en el caso donde las relaciones complejas están involucradas. También la regresión múltiple puede ser el camino correcto cuando se ha determinado que más de una variable afecta las propiedades de la roca.

### **PREDICCIÓN DE LA RESISTENCIA COMPRESIVA (UCS) EN SHALES DE LA FORMACIÓN BREATHITT USANDO DURABILIDAD, DUREZA Y PROPIEDADES ESTRUCTURALES DE LA ROCA<sup>13</sup>**

La resistencia a la compresión es uno de los más amplios parámetros usados en diseño en la industria de la construcción y la ingeniería de rocas. Sin embargo las pruebas estándar establecidas para la preparación de las muestras son muy exigentes. De cualquier forma es bastante difícil y en ocasiones imposible cumplir estos requerimientos usando rocas débiles y especialmente shales.

En el presente estudio se evaluó el uso del ensayo de durabilidad Slake y el uso de ensayo de dureza Shore para estimar el UCS.

El ensayo de Resistencia la compresión simple ha sido estandarizado por la Norma ASTM y ISRM. Este involucra extracción del bloque de roca o del corazón con brocas. Desafortunadamente la existencia del shale en diferentes formas que van desde rocas lodosas suaves hasta metamórficas duras, dificultan la toma de muestras y la preparación.

La inherente debilidad como la fisilidad y la Solubilidad en agua empeora su situación. Por tal razón, la mayoría del tiempo los Shales no son ensayados en detalle y el valor la resistencia compresiva uniaxial encontrado de un pequeño grupo de muestras que suelen ser los más fuertes y los más fáciles de preparar que los otros, se asume como la caracterización de toda la masa rocosa.

En general se han buscado realizar ensayos más simples y utilizar estos resultados para predecir propiedades que involucran preparación de muestras y ensayos más exigentes.

### **Factores que Afectan las propiedades mecánicas de los shale**

---

<sup>13</sup> KONKAGUL, Engin y SANTI, Paul. Predicting the unconfined compressive strength of the Breathitt Shales using Slake durability, shores hardness and rock structural properties. International Journal of Rock Mechanics and Mining Sciences 36 (1999) 139-153

La mineralogía y la disposición geométrica de las partículas (microfabricas) afectan la resistencia de las rocas débiles. La dificultad con los shales es que la microfabrica incluye características de ambos, rocas arcillosas (rocas lodosas y rocas arcillosa) y de las rocas frágiles (Compuestas de granos, como las arenas).

Para poder predecir el comportamiento de los shales es importante entender el efecto de los granos como minerales arcillosos.

**En las rocas tipo shale la mineralogía** puede variar en: cantidades de minerales arcillosos, minerales no arcillosos, materia orgánica y sales precipitadas. La mineralogía es el primer factor que controla propiedades físicas y químicas de tales rocas. Para aplicaciones ingenieriles los minerales arcillosos se refiere a aquellos materiales con tamaño inferior a 0.002 mm. Como termino mineral se refiere a los minerales arcillosos específicos (esmectita, talco, clorita,). Se caracterizan también por tener carga residuales negativas que son balanceadas por la absorción de cationes.

Otro de los factores que afecta estas rocas es el **Contenido de agua ( $w_c$ )**, en general, el agua puede debilitar los enlaces o interactuar con las superficies de los minerales y alterar sus propiedades superficiales. Además a través de la presión de poro del agua, esta puede causar inestabilidad a través de los planos de debilidad.

El agua también puede disminuir la resistencia al corte friccional o cambiar las características de los minerales arcillosos que constituyen la roca.

**La porosidad** es tomada en cuenta por varios autores como un parámetro importante que afecta las propiedades mecánicas de las rocas, por tanto entre más grandes los poros, más baja la resistencia a la compresión<sup>14</sup>. En efecto Se puede inferir que por lo menos todas las propiedades de resistencia en las rocas sedimentarias decrecen con el incremento de la Porosidad. La razón es porque alta porosidad facilita la creación de redes de microfracturas inducidas por los esfuerzos.

**El grado de alineación y la Presencia de minerales blandos** es uno de los factores que notoriamente influyen en las propiedades mecánicas de los shale, los granos perfectamente alineados pueden formar planos de debilidad (Planos de estratificación, esquistosidad y laminación) los cuales afectan el ensayo de compresión uniaxial y los módulos de elasticidad. Por ejemplo, si una roca fuera cargada paralela a la orientación de los granos, la falla debería ser a lo largo de uno de estos planos de debilidad, y la deformación transversal del plano de debilidad es en la mayoría de los casos mayor a la deformación de la roca intacta, afectando la determinación del módulo de Young y relación de poisson.

---

<sup>14</sup> Deere DU, Miller RP. Engineering Classification and Index Properties for Intact Rock. Department of Civil Engineering, University of Illinois, Urbana, IL 1966, pp. 90

En cuanto a los minerales blandos, estos pueden tener múltiples orígenes: depositados originalmente, ser producto de alteraciones o desgaste. Los más comunes son el talco, la mica, la halita y la calcita. La mayoría de estos minerales son fácilmente disueltos al contacto con el agua o las fuerzas de hidratación destruyen su integridad estructural por rompimiento de enlaces. Los minerales blandos también fallan fácilmente incluso bajo poca carga en resistencia compresiva uniaxial.

## 2 REDES NEURONALES APLICADAS A SHALES

Un tema que es necesario entender para empezar a tratar las redes neuronales es el reconocimiento de patrones.

### ¿Qué es un patrón?<sup>15</sup>

Un Patrón es una entidad a la que se le puede dar un nombre y está representado por un conjunto de propiedades medidas y las relaciones entre ellas, el cual matemáticamente se representaría como un vector de características. Por ejemplo, el patrón puede ser la Onda sonora y su vector de características es el conjunto de coeficientes espectrales extraídos de ella. El reconocimiento de patrones puede tener cualquiera de los dos objetivos:

- ✓ Identificar el patrón como miembro de una clase ya definida, es decir una clasificación supervisada.
- ✓ Asignar el patrón a una clase no definida. (clasificación no supervisada, agrupamiento o clustering)

El diseño de un sistema de reconocimiento de patrones se lleva cabo normalmente en 3 fases:

- ✓ Adquisición y proceso de datos.
- ✓ Extracción de características
- ✓ Toma de decisiones o agrupamiento

Tradicionalmente el reconocimiento de patrones se ha abordado desde un punto de vista estadístico

### 2.1 Reconocimiento estadístico de patrones (REP)<sup>16</sup>

En estadística el patrón se considera como un vector numérico de dimensión N. Por tanto un patrón es un punto en el espacio N-dimensional (de características). El REP funciona de dos modos diferentes: Entrenamiento y reconocimiento.

- ✓ **Modo entrenamiento:** Se diseña el vector de características para representar los patrones de entrada y se entrena al clasificador con un conjunto de datos de entrenamiento de forma que el número de patrones mal identificados se minimice.
- ✓ **Modo de reconocimiento:** el clasificador ya entrenado toma como entrada el vector de características de un patrón desconocido y lo asigna a una de las clases o categorías.

---

<sup>15</sup> ALONSO, Luis y CALONGE, Teodoro. Capítulo 1 Redes Neuronales y Reconocimiento de patrones. En ALONSO, L. Reconocimiento de Patrones con Redes Neuronales. Salamanca: Imprenta cardenal, 2001. P. 1.1.

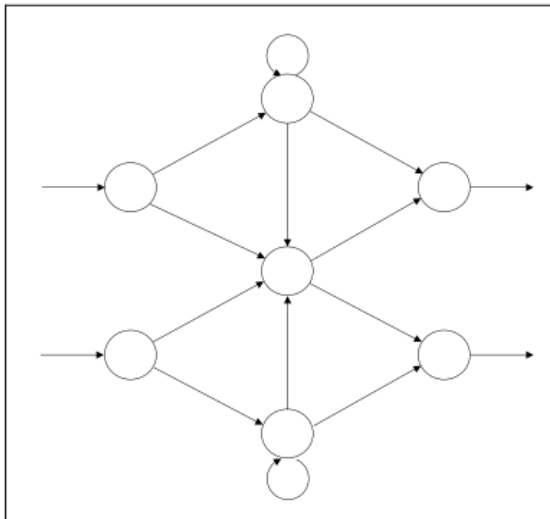
<sup>16</sup> Ibid., p.1.1.

Existe otra alternativa estudiada en los últimos años, conocida como redes neuronales artificiales.

## 2.2 Redes neuronales artificiales<sup>17</sup>

La Neurocomputación es un aporte más al viejo objetivo de crear sistemas inteligentes, considerando como tales a máquinas capaces de llevar a cabo tareas que exhiben alguna de las características asociadas a la inteligencia humana.

Desde el punto de vista Computacional una Red Neuronal puede ser descrita como un conjunto de autómatas celulares (neuronas), figura 2-1, por el cual se establece un flujo de información mediante una topología de interconexiones (sinapsis).



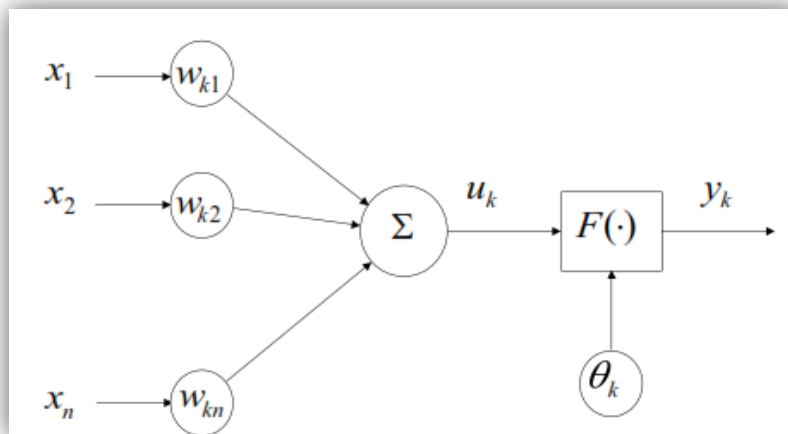
**Figura 2-1.** Autómatas Celulares

**Fuente:** ALONSO, Luis. Redes Neuronales y Reconocimiento de patrones. Salamanca, 2001.

En la mayoría de las RNA propuestas, cada neurona procesa la información según el modelo propuesto por McCulloch-Pitts

---

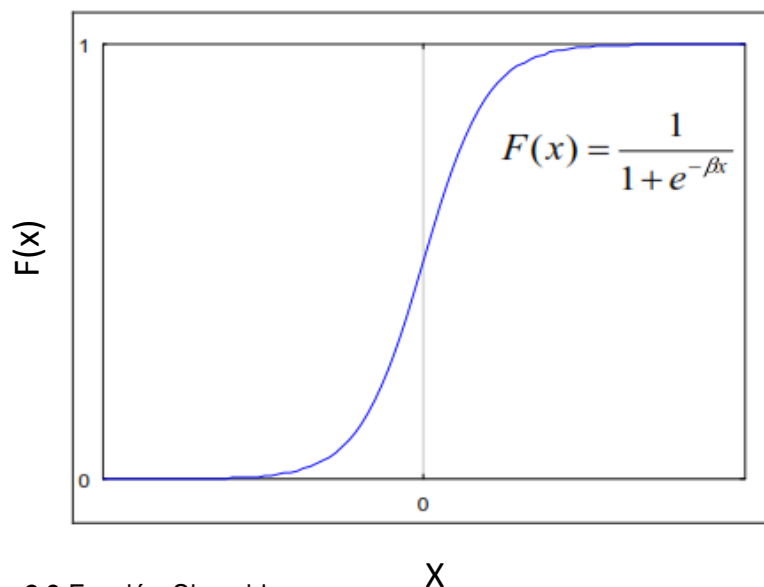
<sup>17</sup> Ibid., p.1.2.



**Figura 2-2.** Neurona artificial (McCulloch-Pitts)

**Fuente:** ALONSO, Luis. Redes Neuronales y Reconocimiento de patrones. Salamanca, 2001.

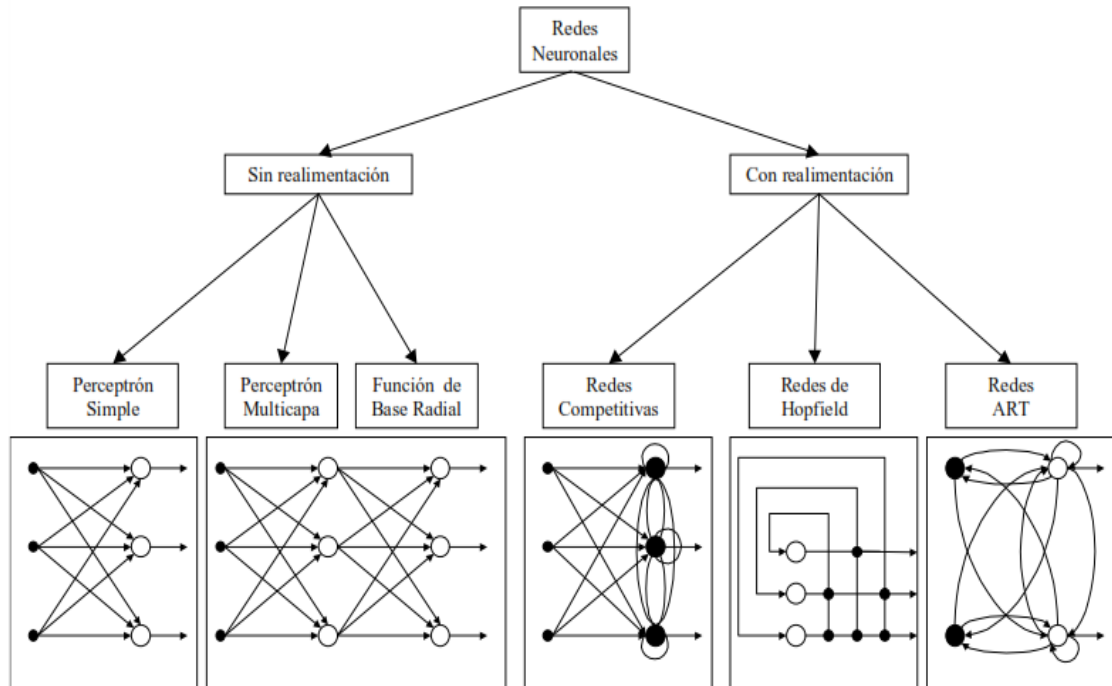
En síntesis cada neurona funciona de la siguiente manera (figura 2-2): Obtiene su salida como la suma ponderada de cada una de sus entradas previa manipulación por su correspondiente peso. Si dicho peso es positivo, se dirá que la entrada es excitadora; en caso contrario se le llamará inhibitoria como en el modelo biológico. La salida de valores es filtrada por una función de activación  $F(X)$ , que bien podría tratarse de una función de salto (0,1) desplazada del origen cierto umbral  $\theta_k$ . Esta función estaría en sintonía con los conocimientos fisiológicos acerca del funcionamiento de las neuronas, puesto que cuando una determinada señal nerviosa (electroquímica) no supera un cierto Umbral, esta se convierte en inhibitoria y viceversa. A pesar de esto, no es la función salto la más comúnmente empleada sino otra: la sigmoide, Figura 2-3. Se usa por conveniencia numérica, ya que es una función  $n$ -diferenciable.



**Figura 2-3** Función Sigmoide

**Fuente:** ALONSO, Luis. Redes Neuronales y Reconocimiento de patrones. Salamanca, 2001.

En la práctica encontrar la red y los pesos adecuados no es un problema fácil, actualmente se trabaja en cierto tipo de redes, los cuales no son universales, sino que trabajan bien para algunas tareas y para otras no. En particular se habla de redes con o sin lazos de realimentación. A continuación según este criterio se clasifican las redes más conocidas:



**Figura 2-4** Taxonomía de RNA

**Fuente:** ALONSO, Luis. Redes Neuronales y Reconocimiento de patrones. Salamanca, 2001.

Asociado a cada tipo de red existe al menos un algoritmo de aprendizaje. Este es un método sistemático para encontrar un valor adecuado de los pesos. En términos generales se basa en definir una función objetivo implícita o explícitamente, que representa de alguna forma el estado de la red. A partir de esta los pesos van evolucionando hasta valores que lleven la función al mínimo. Por esto el aprendizaje es característico del tipo del tipo de red.

Existen dos clases de aprendizaje, que es **el aprendizaje supervisado** y el no supervisado; el primero maneja pares entrada-Salida deseada. De esta manera para una determinada entrada se compara la salida obtenida con la red, con la salida deseada. Por lo tanto el ajuste de los pesos está dirigido enfocado a la disminución de la diferencia entre la salida de la red y la respuesta deseada.

En **el aprendizaje no supervisado**, el planteamiento es diferente. No existen salidas deseadas, aunque a la red se le adicionan entradas para su entrenamiento, y el sistema evoluciona de forma auto-organizada hasta encontrar un estado que se considera estable.

La red aprende con ejemplos, algo realmente atractivo de la red es que tiene una alta capacidad para generalizar, incluso da muy buenas respuestas cuando se introducen entradas que no se utilizan en el entrenamiento.

Por tanto se conocen dos fases en el funcionamiento de una red neuronal, el entrenamiento y el reconocimiento. El principio es fijar los pesos en la fase de entrenamiento y luego pasa al reconocimiento, en esta ya realiza procesos de la aplicación real.

En predicción son muy utilizadas las redes. En este caso se pretende averiguar cuál es el siguiente término en una serie temporal, por tanto la red es entrenada con datos conocidos hasta cierto momento, y se espera que el valor se obtenga para datos posteriores.

### **2.3 Predicción de Propiedades Mecánicas de la Roca Utilizando Redes Neuronales: Revisión de la Literatura.**

#### **Predicción de La Resistencia Compresiva Uniaxial Y Módulos de Elasticidad Para Muestras Travertino Usando Regresión Y Redes Neuronales Artificiales<sup>18</sup>**

Las redes neuronales pueden ser aplicadas en el aprendizaje relacionado con la clasificación, generalización, caracterización y Optimización de las funciones. Porque las redes neuronales tienen la capacidad de trabajar con datos incompletos, poseer un error de tolerancia y mostrar una convergencia gradual. Fácilmente pueden formar modelos para problemas complejos. Especialmente en el desarrollo de soluciones para semi-estructural o problemas no estructurales, los modelos de redes neuronales artificiales pueden proveer exitosos resultados. La mayoría de ellos son fáciles, más rápidos y más adaptables que los métodos tradicionales.

¿Cómo funcionan?: Es un sistema de procesamiento distribuido paralelo compuesto de dos componentes: el nodo (Elemento de procesamiento, neurona artificial o unidad) y la conexión. Un Parámetro  $W_{ij}$  (Peso) está asociado con cada conexión entre dos celdas. Entonces cada celda en la capa superior recibe el peso entrado desde cada nodo en la capa abajo y entonces se procesa esta colección de entrada antes que la unidad envíe una señal a otras capas.

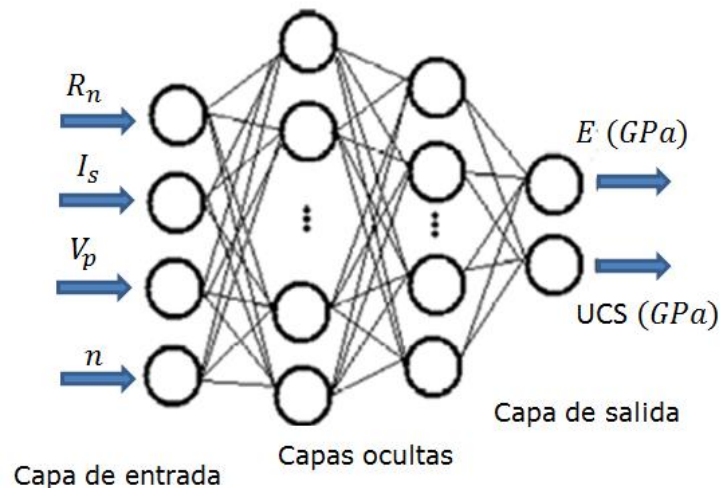
En estudio de DEGHAN, SATTARI y CHEHREH se utilizó una regresión generalizada red neuronal y una red de alimentación de múltiples capas hacia adelante para predecir UCS y E.

En este trabajo se estudió dos clases de Redes neuronales GRNN y FANN

---

<sup>18</sup> DEGHAN, S. SATTARI, GH. CHEHREH CHELGANIS. " Prediction of Uniaxial Compressive Strength and Modulus of Elasticity for Travertine Sample Using Regression and Artificial Neural Networks. Mining Science and Technology 0041-0046, 2009.

GRNN (regresión generalizada red neuronal): Es una base radial de la red neuronal. La principal ventaja de la regresión de la red es su extremadamente rápido entrenamiento, el cual no requiere un procedimiento de entrenamiento iterativo, y aproxima cualquier función arbitraria entre la entrada y los vectores de salida, dibujando la función de estimación directamente desde los datos de entrenamiento. Un esquemático es presentado.



**Figura 2-5** Arquitectura de GRNN

**Fuente:** modificado de: DEGHAN, S. SATTARI, Prediction of Uniaxial Compressive Strength and Modulus of Elasticity for Travertine Sample Using Regression and Artificial Neural Networks. Mining Science and Technology, 2009.

FANN (Alimentador Multicapa Hacia Delante De La Red): Tiene sus neuronas organizadas en capas sin alimentar hacia atrás o conexiones laterales. La configuración FANN consta de una capa de entrada, una o más ocultas y una capa de salida. En un FANN las cantidades son alimentadas en los nodos de entrada, que una vez los transmitirá a los nodos de la capa oculta tras multiplicar por un peso. Un nodo de capa oculta cuya función es intervenir entre la entrada externa y la salida de la red, se le adiciona el peso de entrada recibido desde cada nodo de entrada, se asocia con una tendencia y luego pasan los resultados a los nodos de la siguiente capa oculta o salida, a través de una función de transferencia no lineal. El proceso de aprendizaje trabaja en pequeños pasos iterativos. La salida es comparada a los conocidos y buenos resultados y una señal de error cuadrático es calculada. El valor del error es propagado hacia atrás a través de la red y se realizan cambios pequeños de peso en cada capa. Los cambios de peso se calculan para reducir la señal de error del caso en cuestión. El ciclo se repite hasta que el valor global de error cae por debajo de un umbral predeterminado

## ESTIMACION DE LAS CONSTANTES ELASTICAS DE LAS ROCAS UTILIZANDO UN ENFOQUE ANFIS.

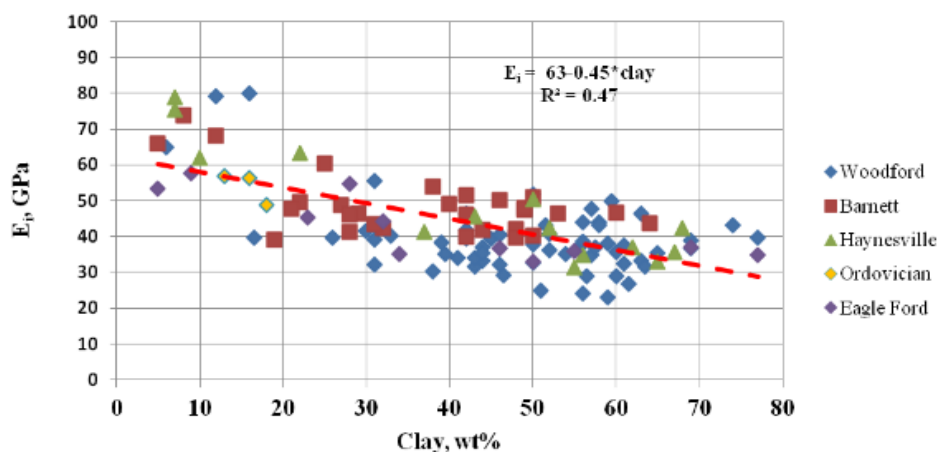
El módulo de Young proporciona una relevante cantidad de magnitudes y características de la deformación de las masas rocosas debido al cambio de esfuerzos en el campo. La determinación del módulo de Young en el laboratorio consume mucho tiempo y es muy costoso. Por lo tanto algunas propiedades básicas de la roca como la carga puntual, densidad y la absorción de agua pueden ser fácilmente determinadas en campo así como en el laboratorio. El trabajo se enfoca en la aplicación de un sistema neuro difuso para la predicción del módulo de Young de la roca. Por tanto las propiedades elásticas (módulo de Young) son determinadas utilizando propiedades físico mecánicas ya mencionadas (carga puntual, densidad y absorción de agua) disminuyendo en gran medida la dificultad para la determinación de la variable.<sup>19</sup>

### 2.4 MODULO DE YOUNG VS PROPIEDADES PETROFISICAS<sup>20</sup>

La arcilla, TOC y la Porosidad son componentes comunes de los shales y todos muestran una correlación negativa con el módulo de Young. Sus efectos en datos combinados desde los ricos en Cuarzo hasta ricos en carbonato fueron investigados.

Existe una total correlación negativa del módulo de Young con el contenido de arcilla independiente del lugar que se encuentre el shale. La razón es porque las muestras ricas en arcilla se comportan dúctilmente y las muestras con deficiencia de arcilla muestra un comportamiento más frágil.

En la Figura 2-6 se observa el efecto del contenido de arcilla en el comportamiento del módulo de Young.



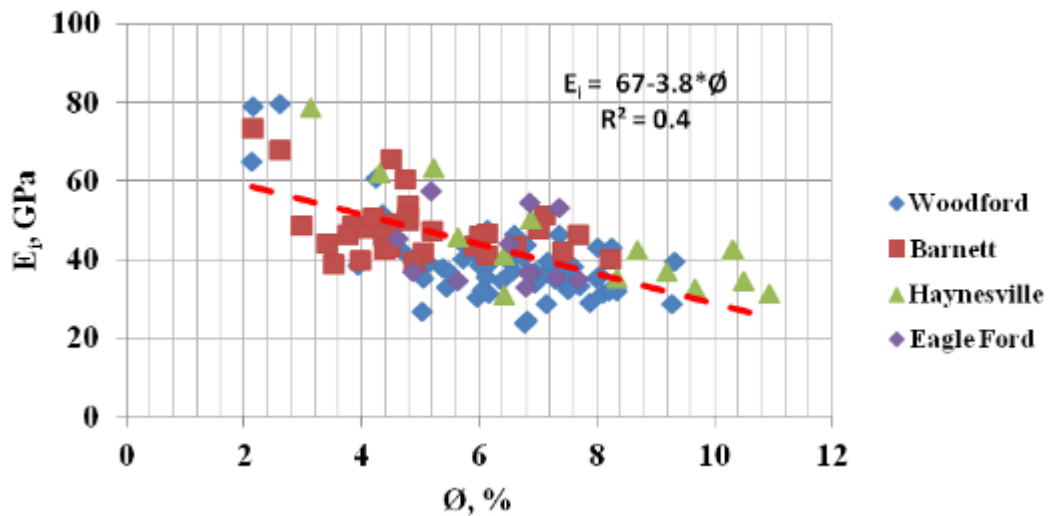
<sup>19</sup> SINGH Rajesh, KAINTHOLA Ashutosh, SINGH T. N; Estimation of elastic constant of rocks using an ANFIS approach, department of Earth sciences , indian institute of technology Bombay; India. Elsevier: Septiembre de 2011.

<sup>20</sup> VIKAS Kumar, Tesis: Master en ciencias : "Geomechanical characterization of shale using Nano-indentation", Universidad de Oklahoma, Norman: Oklahoma 2012.

**Figura 2-6.** Módulo de Young Vs Contenido de arcilla

**Fuente:** “Geomechanical characterization of shale using Nano-indentation”

El efecto de la porosidad debería ser visible en la medida del módulo de Young. En la figura siguiente se observa el módulo de Young versus la porosidad. Lo cual muestra una tendencia general de decrecimiento del módulo de Young con el incremento de la porosidad para todos los shales. El incremento de la porosidad debilita la estructura de la matriz de la muestra haciendo que decrezca el módulo de Young.



**Figura 2-7.** Módulo de Young Vs Porosidad

**Fuente:** Tesis “Geomechanical characterization of shale using Nano-indentation”.

### 3 VALIDACIÓN DEL ANÁLISIS MULTIVARIADO Y LAS REDES NEURONALES.

Para tener certeza en los resultados obtenidos de la técnica desarrollada, se recurre a la comparación con datos de literatura presentados en anteriores estudios.

#### 3.1 Análisis Multivariado

La validación de la técnica de Análisis multivariado se realizó utilizando como apoyo los datos del artículo “Predicción de la resistencia compresiva uniaxial y el módulo de Elasticidad para muestras de Travertino Usando Regresión y Redes Neuronales artificiales”. En este estudio se determina la Resistencia compresiva y el módulo de Young a partir de las variables: Velocidad de Onda compresional ( $V_p$ ), Índice de Carga Puntual ( $I_s$ ), Rebote de Martillo Schmidt ( $R_n$ ) y Porosidad ( $n$ ). Los datos presentados en artículo<sup>21</sup> a partir los ensayos nombrados anteriormente son:

**Tabla 3-1.** Resultados experimentales de las pruebas para la determinación de E y UCS.

N	$I_s$ (Mpa)	$V_p$ (Km/s)	$R_n$	$n$ (%)	E (Gpa)	UCS (Mpa)
1	3	5.66	27.42	3.35	11.45	63.68
2	2.86	5.38	27.38	9.35	8.03	47.46
3	3.63	5.09	30.13	8.14	6.15	37.89
4	3.1	5.81	25.75	2.24	9.35	56.08
5	3.39	5.35	27.38	9.76	5.89	43.46
6	2.54	4.82	27.63	9.64	3.85	25.17
7	3.16	5.59	26	4.5	5.33	30.96
8	4.01	5.25	29.75	8.32	3.69	27.08
9	3.74	5.39	29.13	9.45	4.27	41.51
10	2.73	5.54	26.13	2.11	5.22	28
11	3.08	5.23	26.88	9.22	3.18	30.13
12	3.6	5.4	28	7.42	3.67	32.35
13	3.02	5.73	26.5	3.95	5.34	41.12
14	2.26	5.67	26.75	2.93	5.17	49.63
15	3.36	5.1	29.63	6.38	5.04	41.55
16	3.3	5.52	27.5	9.28	6.19	58.2
17	3.52	5.49	27.38	3.97	5.21	32.92
18	2.63	5.05	30.5	9.51	3.48	25.82
19	2.91	5.42	26.63	4.89	5.04	30.06
20	3.06	5.51	30.3	3.78	6.08	35.38
21	3.89	5.45	25.63	0.97	5.64	39.19

DEGHAN, OP. cit., p. 42-43.

N	I <sub>s</sub> (Mpa)	V <sub>p</sub> (Km/s)	R <sub>n</sub>	n (%)	E (Gpa)	UCS (Mpa)
22	3.32	5.09	27.63	8.98	4.16	35.63
23	3.32	5.02	28.63	7.99	3.05	22.7
24	2.61	4.84	29.88	6.66	3.67	27.99
25	3.26	5.4	29.25	8.93	6.13	49.62
26	3.02	5.79	27.38	3.49	7.15	56.56
27	3.87	5.19	26.25	9.13	3.77	28.69
28	3.17	5.13	28.25	9.29	3.52	29.53
29	2.92	5.09	28.88	10.27	3.41	29.91
30	2.98	5.67	27.63	3.48	8.75	71.54

**Fuente:** Datos tomados de "Prediction of uniaxial compressive strength and modulus of elasticity for Travertine samples using regression" Mining Science and Technology 20 (2010) 0041–0046.

### 3.1.1 Correlaciones para la Resistencia Compresiva Uniaxial (UCS) y el Modulo de Young (E) determinadas en Artículo (DEGHAN, SATTARI, CHEHREH; 2010)

El artículo muestra que se trabajó con una población de 30 datos para cada uno de los ensayos. A estos resultados experimentales se aplicó un análisis multivariado en el cual una serie de ecuaciones fueron ajustadas. Finalmente los autores muestran dos correlaciones de Regresión Multivariada para la predicción del módulo de Elasticidad y la Resistencia compresiva Uniaxial, estas son las siguientes:

$$\begin{aligned}
 UCS = & -595.303 - 442.363V_p + 45.338V_p^2 - 6.1n \\
 & + 0.52n^2 + 28.314I_{s(50)} - 4.06I_{s(50)}^2 \\
 & + 115.822R_n - 2.007R_n^2
 \end{aligned}
 \tag{Ecuación 3-1}$$

$$R^2 = 0.64$$

$$\begin{aligned}
 E = & 32.9 - 53.585V_p + 5.459V_p^2 + 5.693R_n - \\
 & 0.096R_n^2 - 0.898n + 0.065n^2 + 13.162I_{s(50)} - \\
 & 2.038I_{s(50)}^2
 \end{aligned}
 \tag{Ecuación 3-2}$$

$$R^2 = 0.56$$

A partir de los datos del Artículo mencionado anteriormente, Tabla 3-1, se trabajó el análisis multivariado utilizando la herramienta Excel. Como en las ecuaciones anteriores se observa que las variables utilizadas aparecen también elevadas al cuadrado y son datos de entrada del análisis, estas se organizaron en la Tabla 3-2.

**Tabla 3-2** Variables independientes modelo de Regresión multivariado

N°	Vp (Km/s)	Vp <sup>2</sup>	n (%)	n <sup>2</sup>	Is (Mpa)	Is <sup>2</sup>	Rn	Rn <sup>2</sup>
1	5.66	32.0356	3.35	11.2225	3	9	27.42	751.8564
2	5.38	28.9444	9.35	87.4225	2.86	8.1796	27.38	749.6644
3	5.09	25.9081	8.14	66.2596	3.63	13.1769	30.13	907.8169
4	5.81	33.7561	2.24	5.0176	3.1	9.61	25.75	663.0625
5	5.35	28.6225	9.76	95.2576	3.39	11.4921	27.38	749.6644
6	4.82	23.2324	9.64	92.9296	2.54	6.4516	27.63	763.4169
7	5.59	31.2481	4.5	20.25	3.16	9.9856	26	676
8	5.25	27.5625	8.32	69.2224	4.01	16.0801	29.75	885.0625
9	5.39	29.0521	9.45	89.3025	3.74	13.9876	29.13	848.5569
10	5.54	30.6916	2.11	4.4521	2.73	7.4529	26.13	682.7769
11	5.23	27.3529	9.22	85.0084	3.08	9.4864	26.88	722.5344
12	5.4	29.16	7.42	55.0564	3.6	12.96	28	784
13	5.73	32.8329	3.95	15.6025	3.02	9.1204	26.5	702.25
14	5.67	32.1489	2.93	8.5849	2.26	5.1076	26.75	715.5625
15	5.1	26.01	6.38	40.7044	3.36	11.2896	29.63	877.9369
16	5.52	30.4704	9.28	86.1184	3.3	10.89	27.5	756.25
17	5.49	30.1401	3.97	15.7609	3.52	12.3904	27.38	749.6644
18	5.05	25.5025	9.51	90.4401	2.63	6.9169	30.5	930.25
19	5.42	29.3764	4.89	23.9121	2.91	8.4681	26.63	709.1569
20	5.51	30.3601	3.78	14.2884	3.06	9.3636	30.3	918.09
21	5.45	29.7025	0.97	0.9409	3.89	15.1321	25.63	656.8969
22	5.09	25.9081	8.98	80.6404	3.32	11.0224	27.63	763.4169
23	5.02	25.2004	7.99	63.8401	3.32	11.0224	28.63	819.6769
24	4.84	23.4256	6.66	44.3556	2.61	6.8121	29.88	892.8144
25	5.4	29.16	8.93	79.7449	3.26	10.6276	29.25	855.5625
26	5.79	33.5241	3.49	12.1801	3.02	9.1204	27.38	749.6644
27	5.19	26.9361	9.13	83.3569	3.87	14.9769	26.25	689.0625
28	5.13	26.3169	9.29	86.3041	3.17	10.0489	28.25	798.0625
29	5.09	25.9081	10.27	105.4729	2.92	8.5264	28.88	834.0544
30	5.67	32.1489	3.48	12.1104	2.98	8.8804	27.63	763.4169

Los datos anteriores se usaron como entrada para determinar la correlación, estos mismos corresponden a las variables independientes, el UCS y el E se introdujeron como las variables dependientes en cada caso.

### 3.1.2 Correlación para el cálculo del UCS:

Para cada una de las variables independientes se obtiene los coeficientes que minimizan el error para determinar el UCS

**Tabla 3-3** Estadística de la Regresión de UCS para cada Variable.

	<i>Coeficientes</i>	<i>Error típico</i>	<i>Estadístico t</i>	<i>Probabilidad</i>	<i>Inferior 95%</i>	<i>Superior 95%</i>
intercepción	-573.408	1133.527	-0.506	0.618	-2930.706	1783.890
Vp	-447.734	285.125	-1.570	0.131	-1040.684	145.215
Vp <sup>2</sup>	45.862	26.950	1.702	0.104	-10.183	101.906
n	-6.009	4.436	-1.354	0.190	-15.235	3.217
n <sup>2</sup>	0.514	0.337	1.525	0.142	-0.187	1.214
Is	27.952	51.477	0.543	0.593	-79.100	135.003
Is <sup>2</sup>	-4.007	8.019	-0.500	0.623	-20.683	12.669
Rn	115.293	59.746	1.930	0.067	-8.956	239.542
Rn <sup>2</sup>	-1.998	1.056	-1.892	0.072	-4.195	0.198

Si se extraen los coeficientes y se arma la ecuación para obtener la Resistencia compresiva uniaxial, se tiene:

$$\begin{aligned}
 UCS = & -573.408 - 447.734V_p + 45.862V_p^2 - \\
 & 6.009n + 0.514n^2 + 27.952I_{s(50)} - 4.007I_{s(50)}^2 + \\
 & 115.293R_n - 1.998R_n^2
 \end{aligned}
 \tag{Ecuación 3-3}$$

El coeficiente de determinación  $R^2$ , para la ecuación anterior es:

$$R^2 = 0.646$$

Este valor corresponde al mismo que se reporta en el artículo.

### 3.1.3 Correlación para Cálculo de E

De igual manera se utilizaron los datos de entrada que se encuentran en la Tabla 3-2 como variables independientes y el módulo de Young como variable dependiente y finalmente se obtuvieron los siguientes resultados.

**Tabla 3-4.** Estadística de la Regresión del Módulo de Young para cada Variable.

	<i>Coeficientes</i>	<i>Error típico</i>	<i>Estadístico t</i>	<i>Probabilidad</i>	<i>Inferior 95%</i>	<i>Superior 95%</i>
intercepción	36.107	199.358	0.181	0.8580	-378.480	450.693
Vp	-54.234	50.146	-1.082	0.2917	-158.519	50.050
Vp <sup>2</sup>	5.523	4.740	1.165	0.2570	-4.334	15.380
n	-0.884	0.780	-1.133	0.2698	-2.507	0.738
n <sup>2</sup>	0.064	0.059	1.076	0.2939	-0.059	0.187

	<i>Coefficien- tes</i>	<i>Error típico</i>	<i>Estadís- tico t</i>	<i>Probabi- lidad</i>	<i>Inferior 95%</i>	<i>Superior 95%</i>
Is	13.105	9.053	1.447	0.1625	-5.723	31.932
Is^2	-2.029	1.410	-1.439	0.1649	-4.962	0.904
Rn	5.588	10.508	0.532	0.6005	-16.264	27.440
Rn^2	-0.094	0.186	-0.505	0.6186	-0.480	0.293

Si se extraen los coeficientes de la Tabla 3-4 y se organiza la correlación obtenida para el cálculo del Módulo de Elasticidad se llega a la siguiente Expresión:

$$\begin{aligned}
 E = & 36.107 - 54.234V_p + 5.523V_p^2 + 5.588R_n - 0.094R_n^2 \\
 & - 0.884n + 0.064n^2 + 13.105I_{s(50)} \\
 & - 2.029I_{s(50)}^2
 \end{aligned}
 \tag{Ecuación 3-4}$$

Si se compara la correlación anterior (Ecuación 3-4), con la ecuación 3-2, se observa que los coeficientes de las variables presentan valores relativamente cercanos aunque no los mismos. El coeficiente de determinación  $R^2$  obtenido fue:

$$R^2 = 0.561$$

Este es el mismo valor que se reporta en el artículo.

Debido a que en la determinación de la ecuación del módulo de Young y UCS en el artículo presenta valores de coeficientes diferentes a los determinados en el presente estudio, se planteó un análisis de varianza con el objeto de determinar si sus resultados estadísticamente son similares o distintos.

#### 3.1.4 Análisis de varianza (ANOVA) para comparación de resultados obtenidos en este estudio con el trabajo de DEGHAN, SATTARI, CHEHREH (2010).

En la Tabla 3-5 se encuentran los datos observados de los ensayos y los resultados usando la ecuación obtenida en el artículo y la ecuación que se determinó en el presente estudio.

**Tabla 3-5.** Datos Observados y resultados de las ecuaciones.

N	Observados		Artículo		Presente Estudio	
	E (Gpa)	UCS (Mpa)	E (Gpa)	UCS (Mpa)	E (Gpa)	UCS (Mpa)
1	11.45	63.68	7.2793485	54.018598	7.4344772	54.190435
2	8.03	47.46	4.78537694	39.8887404	4.9325192	40.0420994
3	6.15	37.89	3.88370961	29.5017075	4.058102	29.5932421
4	9.35	56.08	8.3171329	54.3540463	8.4733017	54.6384096

N	Observados		Articulo		Presente Estudio	
	E (Gpa)	UCS (Mpa)	E (Gpa)	UCS (Mpa)	E (Gpa)	UCS (Mpa)
5	5.89	43.46	5.00143009	41.6962502	5.1452794	41.8160655
6	3.85	25.17	3.12171724	29.0454889	3.2690546	29.2226458
7	5.33	30.96	5.58074654	35.2648918	5.7352139	35.513323
8	3.69	27.08	3.47805169	34.8015195	3.6531932	34.9117429
9	4.27	41.51	5.08420644	46.2627055	5.2479993	46.3701558
10	5.22	28	5.93316301	38.0799005	6.0681685	38.1962221
11	3.18	30.13	4.08518286	31.0878134	4.2260051	31.2254814
12	3.67	32.35	4.75037	40.201008	4.9085696	40.3450096
13	5.34	41.12	7.16926636	50.9000362	7.3313551	51.181372
14	5.17	49.63	5.43084404	46.0141127	5.5880129	46.2585822
15	5.04	41.55	4.13826984	29.2211337	4.3045646	29.2988182
16	6.19	58.2	5.9101606	49.0278532	6.0645868	49.2105124
17	5.21	32.92	5.69647776	42.584817	5.8457947	42.6985548
18	3.48	25.82	3.70375011	27.956503	3.8791938	28.0663181
19	5.04	30.06	4.56624709	30.6372709	4.7136581	30.8067333
20	6.08	35.38	6.46976874	43.5396458	6.6479255	43.6300386
21	5.64	39.19	5.40673059	33.8737787	5.5288956	33.8911867
22	4.16	35.63	4.00490806	32.0814535	4.1484437	32.1840748
23	3.05	22.7	3.98156426	31.1721849	4.1356974	31.2679932
24	3.67	27.99	3.19712099	23.294118	3.3725327	23.4534497
25	6.13	49.62	5.52397694	44.8223745	5.6886582	44.9379304
26	7.15	56.56	8.37610876	63.484893	8.5429055	63.7460416
27	3.77	28.69	2.76124851	23.8625483	2.9058868	24.0354115
28	3.52	29.53	4.39638051	35.9564627	4.542678	36.0392179
29	3.41	29.91	4.61794816	38.941551	4.7648227	38.9995668
30	8.75	71.54	7.36921546	55.4404639	7.5271601	55.6183184

A la data anterior se realizó prueba F para varianza de dos muestras obteniendo los siguientes resultados

### 3.1.4.1 ANOVA Datos observados y Regresión Articulo

✓ Módulo de Young

**Tabla 3-6.** ANOVA módulo de Young de datos observados y regresión artículo de DEHGHAN, SATTARI, CHEHREH (2010)

	<i>Variable 1</i>	<i>Variable 2</i>
Media	5.36266667	5.13401409
Varianza	3.95342713	2.21166277
Observaciones	30	30
Grados de libertad	29	29
F	1.78753613	
P(F<=f) una cola	0.06178864	
Valor crítico para F (una cola)	1.86081144	

✓ Resistencia compresiva uniaxial (UCS)

**Tabla 3-7.** ANOVA UCS de datos observados y regresión artículo de DEHGHAN, SATTARI, CHEHREH (2010)

	<i>Variable 1</i>	<i>Variable 2</i>
Media	38.9936667	39.2337957
Varianza	158.446817	101.475839
Observaciones	30	30
Grados de libertad	29	29
F	1.56142407	
P(F<=f) una cola	0.1180701	
Valor crítico para F (una cola)	1.86081144	

Si se observa los resultados, el valor de F es menor que el valor crítico de F en los dos casos, módulo de Young y UCS. Esto quiere decir que la variación entre las dos muestras es mínima, por tanto se concluye que estadísticamente los resultados son iguales.

### 3.1.4.2 ANOVA Datos observados y Regresión estudio

✓ Módulo de Young

**Tabla 3-8.** ANOVA módulo de Young de datos observados y regresión estudio

	<i>Variable 1</i>	<i>Variable 2</i>
Media	5.36266667	5.28948853
Varianza	3.95342713	2.21311021
Observaciones	30	30
Grados de libertad	29	29
F	1.78636704	
P(F<=f) una cola	0.06199749	
Valor crítico para F (una cola)	1.86081144	

✓ Resistencia compresiva uniaxial (UCS)

**Tabla 3-9.** ANOVA UCS de datos observados y regresión estudio

	<i>Variable 1</i>	<i>Variable 2</i>
Media	38.9936667	39.3796318
Varianza	158.446817	102.135891
Observaciones	30	30
Grados de libertad	29	29
F	1.55133338	
P(F<=f) una cola	0.12148645	
Valor crítico para F (una cola)	1.86081144	

Los resultados del análisis de varianza entre los datos obtenidos a partir de la ecuación del presente estudio y los datos observados evidencian que existe similitud entre los resultados, y se cumple para los dos casos módulo de Young y UCS.

**3.1.4.3 ANOVA de los resultados de la regresión en estudio y el trabajo de DEGHAN, SATTARI Y CHEHREH (2010)**

✓ Módulo de Young

**Tabla 3-10.** ANOVA módulo de Young de Regresión Artículo (DEGHAN, SATTARI, CHEHREH ;2010) y regresión estudio

	<i>Variable 1</i>	<i>Variable 2</i>
Media	5.28948853	5.13401409
Varianza	2.21311021	2.21166277
Observaciones	30	30
Grados de libertad	29	29
F	1.00065446	
P(F<=f) una cola	0.49930325	
Valor crítico para F (una cola)	1.86081144	

✓ Resistencia compresiva uniaxial (UCS)

**Tabla 3-11.** ANOVA UCS de Regresión Artículo (DEHGHAN, SATTARI, CHEHREH ;2010) y regresión estudio

	<i>Variable 1</i>	<i>Variable 2</i>
Media	39.3796318	39.2337957
Varianza	102.135891	101.475839
Observaciones	30	30
Grados de libertad	29	29
F	1.00650452	
P(F<=f) una cola	0.49309568	
Valor crítico para F (una cola)	1.86081144	

De igual forma que en el caso anterior, el valor de F es menor que el valor crítico de F, demostrando que existe similitud entre los datos que resultaron de las dos regresiones.

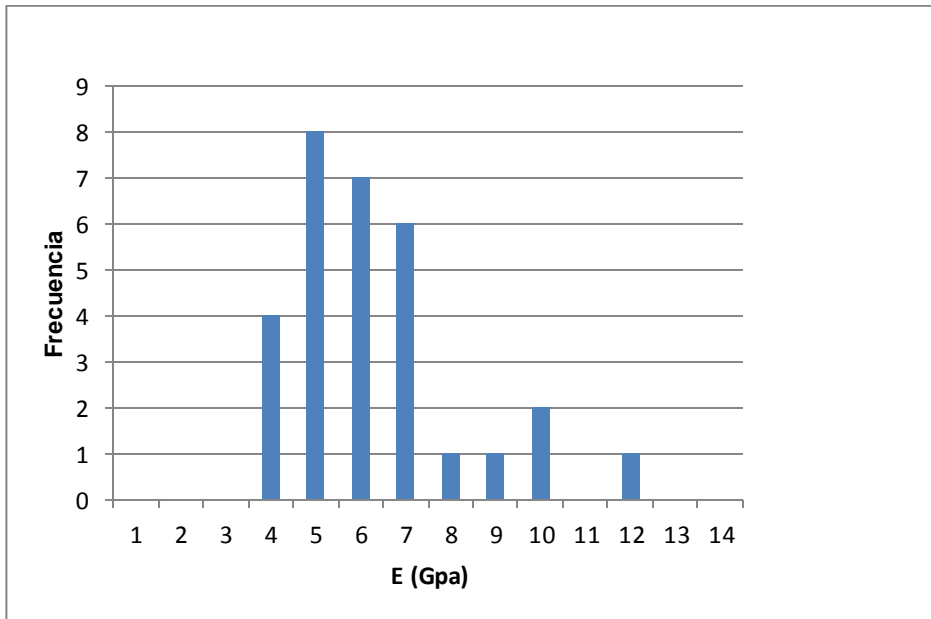
En general, a partir de estos resultados se demuestra que así los coeficientes en las dos regresiones no tengan el mismo valor, las predicciones a partir de las mismas son estadísticamente similares.

### **3.1.5 Tratamiento Estadístico para mejorar las correlaciones en el Cálculo del UCS y Young.**

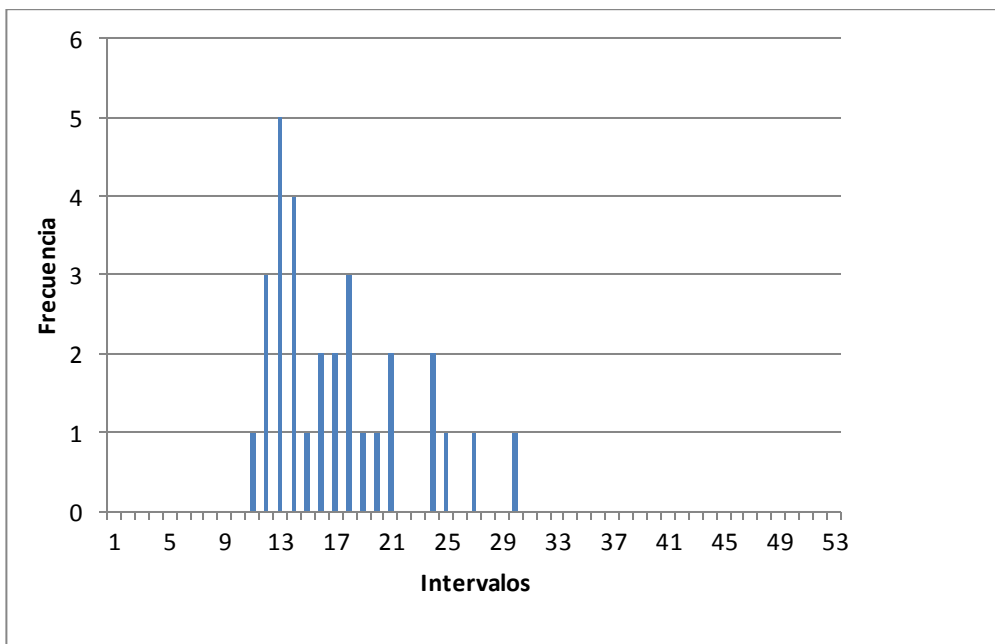
Se realizó un estudio estadístico con el objeto de mejorar la correlación e incrementar el coeficiente de determinación que representa la muestra. La técnica que se decidió utilizar fue un Test de Grubbs con el objeto de descartar datos no representativos.

#### **3.1.5.1 Distribución de los Datos**

Se representaron los resultados de la variable dependiente en este caso el módulo de Young y la resistencia compresiva uniaxial en un gráfico de frecuencias para observar el tipo de distribución que presentaban los datos.



**Figura 3-1** Curva de frecuencias módulo de Young



**Figura 3-2.** Curva de frecuencias Resistencia Compresiva Uniaxial (UCS)

En los dos casos se puede observar una distribución normal con sesgo a la derecha o con sesgo positivo.

### 3.1.5.2 El test de Grubbs

El test de Grubbs utiliza una estadística de prueba, T, que es la diferencia absoluta entre el valor atípico,  $X_0$ , y el promedio de la muestra ( $\bar{X}$ ), dividida en la desviación estándar de la muestra, S.<sup>22</sup>

$$T = \frac{|X_0 - \bar{X}|}{s} \quad \text{Ecuación 3-2}$$

✓ Para un n igual a 30 datos, y un riesgo del 5%, el valor tabulado es de T= 1.697.

✓ Para un n igual a 30 datos y un riesgo del 10%, el valor tabulado es de T=1.310

En la Tabla 3-12 se observan los resultados de T calculados a partir de la ecuación 3-2.

**Tabla 3-12** Resultado de Cálculo de T

N	Vp	Vp	n	n^2	ls	ls^2	Rn	Rn^2	E	UCS
1	1.119	1.131	1.106	1.128	0.418	0.467	0.316	0.335	3.062	1.961
2	0.090	0.065	0.949	1.005	0.752	0.772	0.344	0.362	1.341	0.673
3	0.977	0.981	0.535	0.412	1.085	1.089	1.574	1.600	0.396	0.088
4	1.671	1.724	1.486	1.302	0.180	0.239	1.480	1.437	2.005	1.357
5	0.021	0.046	1.090	1.224	0.512	0.462	0.344	0.362	0.265	0.355
6	1.970	1.904	1.048	1.159	1.515	1.416	0.169	0.192	0.761	1.098
7	0.862	0.860	0.712	0.875	0.037	0.099	1.306	1.277	0.016	0.638
8	0.389	0.411	0.596	0.495	1.991	2.171	1.309	1.318	0.841	0.946
9	0.126	0.103	0.983	1.057	1.347	1.391	0.877	0.865	0.550	0.200
10	0.678	0.668	1.531	1.318	1.062	1.043	1.215	1.193	0.072	0.873
11	0.462	0.483	0.905	0.937	0.227	0.285	0.692	0.699	1.098	0.704
12	0.163	0.140	0.288	0.099	1.013	1.009	0.089	0.064	0.851	0.528
13	1.377	1.406	0.901	1.005	0.371	0.422	0.957	0.951	0.011	0.169
14	1.156	1.170	1.250	1.202	2.183	1.917	0.783	0.786	0.097	0.845
15	0.940	0.946	0.068	0.303	0.440	0.386	1.225	1.229	0.162	0.203
16	0.604	0.592	0.925	0.968	0.297	0.237	0.260	0.281	0.416	1.526
17	0.494	0.478	0.894	1.001	0.822	0.796	0.344	0.362	0.077	0.483
18	1.124	1.121	1.004	1.089	1.301	1.243	1.832	1.879	0.947	1.047
19	0.237	0.214	0.579	0.773	0.633	0.665	0.867	0.865	0.162	0.710
20	0.568	0.554	0.959	1.042	0.275	0.331	1.693	1.728	0.361	0.287
21	0.347	0.327	1.922	1.416	1.705	1.818	1.564	1.514	0.139	0.016
22	0.977	0.981	0.822	0.815	0.345	0.287	0.169	0.192	0.605	0.267
23	1.235	1.225	0.483	0.345	0.345	0.287	0.528	0.506	1.163	1.294
24	1.897	1.837	0.028	0.201	1.349	1.282	1.400	1.414	0.851	0.874

<sup>22</sup> THOMAS, M. y LAU T, Alex. Dealing with Outliers. ASTM standardization, noviembre 2008.

N	Vp	Vp	n	n <sup>2</sup>	Is	Is <sup>2</sup>	Rn	Rn <sup>2</sup>	E	UCS
25	0.163	0.140	0.805	0.790	0.202	0.140	0.960	0.952	0.386	0.844
26	1.598	1.644	1.058	1.101	0.371	0.422	0.344	0.362	0.899	1.396
27	0.609	0.627	0.874	0.891	1.657	1.760	1.132	1.115	0.801	0.819
28	0.830	0.840	0.929	0.973	0.013	0.076	0.263	0.238	0.927	0.752
29	0.977	0.981	1.264	1.510	0.609	0.643	0.702	0.685	0.982	0.722
30	1.156	1.170	1.062	1.103	0.466	0.511	0.169	0.192	1.704	2.586

### 3.1.5.3 Determinación de la correlación de la Resistencia Compresiva uniaxial después de aplicar test de grubbs:

Con el objeto de lograr un mejor ajuste en la correlación de resistencia compresiva, no se tuvieron en cuenta los datos con T calculado mayor a 1.697 que corresponde a un riesgo del 5% en todas las variables. Finalmente permanecieron 19 observaciones que tendrían un 95% de probabilidad.

Nuevamente se aplicó análisis multivariado a los 19 datos restantes y los resultados fueron los siguientes:

**Tabla 3-13.** Estadística de la Regresión para UCS aplicando Test de Grubbs

<i>Estadísticas de la regresión</i>	
Coefficiente de correlación múltiple	0.85312378
Coefficiente de determinación R <sup>2</sup>	0.72782018
R <sup>2</sup> ajustado	0.51007633
Error típico	6.90227293
Observaciones	19

**Tabla 3-14.** Análisis de varianza UCS aplicando Test de Grubbs

	<i>Coefficientes</i>	<i>Error típico</i>	<i>Estadístico t</i>	<i>Probabilidad</i>	<i>Inferior 95%</i>	<i>Superior 95%</i>
Intercepción	-764.752	1818.544	-0.420	0.682	-4816.723	3287.217
Vp (Km/s)	70.014	444.481	0.157	0.877	-920.353	1060.381
Vp <sup>2</sup>	-1.563	41.410	-0.037	0.970	-93.830	90.703
n (%)	1.226	7.032	0.174	0.865	-14.443	16.895
n <sup>2</sup>	0.055	0.505	0.110	0.914	-1.070	1.181
Is (Mpa)	184.311	181.225	1.017	0.333	-219.482	588.106
Is <sup>2</sup>	-29.395	27.785	-1.057	0.314	-91.306	32.514
Rn	6.416	79.486	0.080	0.937	-170.691	183.524
Rn <sup>2</sup>	-0.0058	1.409	-0.004	0.996	-3.147	3.135

A partir los resultados que se observan en la Tabla 3-13 y Tabla 3-14, se concluye que las observaciones están mejor representadas por la correlación después de realizar un control de calidad a los datos a través del Test de Grubbs.

### 3.1.5.4 Determinación de la correlación de Modulo de Young Aplicando test de grubbs:

Igualmente en el ejercicio de mejorar la correlación que describe los datos observados en el módulo de Young, se realizó el siguiente procedimiento: en principio se realizó Test de Grubbs con un riesgo del 5% incluyendo los datos de la variable dependiente, pero el coeficiente de determinación presentó una disminución, pasó de 0.561 a 0.482, esto corresponde a un 14%. Luego se aplicó test de Grubbs de 10 % de riesgo a la variable dependiente y 5% en las variables independientes. Los resultados obtenidos fueron los siguientes:

**Tabla 3-15.** Estadística de la Regresión para (E) aplicando Test de Grubbs

<i>Estadísticas de la regresión</i>	
Coefficiente de correlación múltiple	0.84228937
Coefficiente de determinación R <sup>2</sup>	0.70945138
R <sup>2</sup> ajustado	0.45118594
Error típico	0.08255203
Observaciones	18

**Tabla 3-16.** Análisis de varianza (E) aplicando Test de Grubbs

	<i>Coefficientes</i>	<i>Error típico</i>	<i>Estadístico t</i>	<i>Probabilidad</i>	<i>Inferior 95%</i>	<i>Superior 95%</i>
Intercepción	2.287	21.834	0.104	0.918	-47.104	51.680
Vp (Km/s)	4.436	5.330	0.832	0.426	-7.621	16.494
Vp <sup>2</sup>	-0.376	0.496	-0.759	0.467	-1.499	0.7461
n (%)	-0.128	0.096	-1.326	0.217	-0.347	0.0906
n <sup>2</sup>	0.0088	0.0068	1.294	0.227	-0.0066	0.0244
Is (Mpa)	4.794	2.475	1.936	0.084	-0.804	10.393
Is <sup>2</sup>	-0.726	0.374	-1.936	0.084	-1.574	0.1219
Rn	-1.630	0.969	-1.681	0.126	-3.82	0.5627
Rn <sup>2</sup>	0.030	0.0171	1.751	0.113	-0.008	0.0690

Sin contar con un solo dato adicional al 5% de riesgo (19 datos) que se identificó bastante alejado de la media en la variable a predecir (módulo de Young), el cual se encontraba fuera del 90% de probabilidad; la ecuación paso a tener un coeficiente de determinación de 0.7094 mejorando en un 20.1% con respecto al valor inicial 0.561.

### 3.2 Redes Neuronales en el cálculo del UCS y Modulo de Young a partir de otros ensayos (porosidad, carga puntual, velocidad de Onda compresional y Numero de rebotes del martillo Schmidt.

Se realizó el mismo ejercicio de predecir el Modulo de Young y la Resistencia Compresiva uniaxial pero esta vez se utilizó la técnica de redes neuronales artificiales.

La información que presenta el artículo acerca de las redes neuronales no es completa y en principio se puede prever que tratar de obtener el mismo coeficiente de determinación es una lotería. Por ejemplo, el artículo solamente dice la cantidad de datos que utiliza para entrenamiento y para prueba, pero de los 30 datos no dice cuales utiliza para entrenamiento y cuales para pruebas, ni la cantidad de capas Ocultas y neuronas utilizadas en cada capa. Lo que si se puede llegar a comprobar es que las redes son una buena técnica de predicción, que es lo que se ha encontrado en la literatura de los últimos años.

Se utilizó una red tipo perceptrón multicapa, y un algoritmo genético. Este fue desarrollado utilizando Matlab. Se ampliará la información acerca de la herramienta en el capítulo 5.

#### 3.2.1 UCS utilizando redes neuronales

Se realizaron varias corridas con diferentes configuraciones de redes y se determinó para cada una de ellas el coeficiente de determinación. Algunos de los datos obtenidos fueron los siguientes:

**Tabla 3-17.** Coeficiente de determinación para diferentes Configuración de redes neuronales

RED	Neuronas y Capas Ocultas	R <sup>2</sup>
1	7 4 9	0.201
2	6 4 5	0.555
3	5 8 6 1	0.201
4	8 7 10	0.683
5	2 8 8	0.232
6	3 3 9	0.615
7	4 3 9	0.225
8	8 7 4	0.88

En cada una de las corridas se varió los datos de entrenamiento y prueba. Los datos con los cuales se obtuvo un R<sup>2</sup> de 0.880 fueron:

**Tabla 3-18.** Datos utilizados en red neuronal mejor coeficiente de Determinación

Vp (Km/s)	n (%)	Is (Mpa)	Rn	UCS (Mpa)
<b>Datos Entrenamiento</b>				
5.66	3.35	3	27.42	63.68
5.81	2.24	3.1	25.75	56.08
4.82	9.64	2.54	27.63	25.17
5.25	8.32	4.01	29.75	27.08
5.67	2.93	2.26	26.75	49.63
5.05	9.51	2.63	30.5	25.82
5.45	0.97	3.89	25.63	39.19
5.02	7.99	3.32	28.63	22.7
5.09	10.27	2.92	28.88	29.91
5.67	3.48	2.98	27.63	71.54
5.09	8.14	3.63	30.13	37.89
5.35	9.76	3.39	27.38	43.46
5.59	4.5	3.16	26	30.96
5.54	2.11	2.73	26.13	28
5.23	9.22	3.08	26.88	30.13
5.73	3.95	3.02	26.5	41.12
5.1	6.38	3.36	29.63	41.55
5.49	3.97	3.52	27.38	32.92
5.51	3.78	3.06	30.3	35.38
5.09	8.98	3.32	27.63	35.63
4.84	6.66	2.61	29.88	27.99
5.79	3.49	3.02	27.38	56.56
5.13	9.29	3.17	28.25	29.53
<b>Datos Prueba</b>				
5.38	9.35	2.86	27.38	47.46
5.39	9.45	3.74	29.13	41.51
5.4	7.42	3.6	28	32.35
5.52	9.28	3.3	27.5	58.2
5.42	4.89	2.91	26.63	30.06
5.4	8.93	3.26	29.25	49.62
5.19	9.13	3.87	26.25	28.69

### 3.2.2 Módulo de Young Utilizando Redes Neuronales

Se realizaron varias corridas para determinar un coeficiente de determinación aceptable.

**Tabla 3-19.** Coeficiente de determinación y Configuración de redes

RED	Neuronas y Capas Ocultas	R <sup>2</sup>
1	7 9 9 2	0.298
2	3 5	0.599
3	3 3 9	0.349
4	2 8 8	0.395
5	5 10	0.374
6	9 1 4 5	0.55
7	6	0.381
8	8 7 10	0.416
9	3 3 4 5	0.463
10	8 7 10	0.7595
11	2 8 8	0.352
12	3 3 9	0.593
13	3 5	0.333

Los datos que se utilizaron para la configuración de red neuronal con coeficiente de determinación de 0.759 es:

**Tabla 3-20.** Datos entrenamiento y red con mejor ajuste en el módulo de Young

Vp (Km/s)	n (%)	Is (Mpa)	Rn	E (Gpa)
Entrenamiento				
5.66	3.35	3	27.42	11.45
5.81	2.24	3.1	25.75	9.35
4.82	9.64	2.54	27.63	3.85
5.25	8.32	4.01	29.75	3.69
5.67	2.93	2.26	26.75	5.17
5.05	9.51	2.63	30.5	3.48
5.45	0.97	3.89	25.63	5.64
5.02	7.99	3.32	28.63	3.05
5.09	10.27	2.92	28.88	3.41
5.67	3.48	2.98	27.63	8.75
5.38	9.35	2.86	27.38	8.03
5.35	9.76	3.39	27.38	5.89
5.59	4.5	3.16	26	5.33
5.54	2.11	2.73	26.13	5.22
5.23	9.22	3.08	26.88	3.18
5.73	3.95	3.02	26.5	5.34
5.52	9.28	3.3	27.5	6.19
5.49	3.97	3.52	27.38	5.21
5.51	3.78	3.06	30.3	6.08

5.09	8.98	3.32	27.63	4.16
4.84	6.66	2.61	29.88	3.67
5.79	3.49	3.02	27.38	7.15
5.13	9.29	3.17	28.25	3.52
Prueba				
5.59	4.5	3.16	26	5.33
5.39	9.45	3.74	29.13	4.27
5.4	7.42	3.6	28	3.67
5.1	6.38	3.36	29.63	5.04
5.42	4.89	2.91	26.63	5.04
5.4	8.93	3.26	29.25	6.13
5.19	9.13	3.87	26.25	3.77

#### 4 APLICACIÓN DEL ANÁLISIS MULTIVARIADO A SHALES PARA EL CÁLCULO DEL MÓDULO DE YOUNG (E)

Para este ejercicio se utilizaron los resultados de los análisis obtenidos en el muestreo (corazonamiento) realizado al pozo La Luna en la formación La Luna, tablazo y Simití. Se emplearon datos pertenecientes a los 3 miembros de la formación La Luna: Galembo, Pujamana y Salada.

Las muestras para estos ensayos fueron extraídas horizontales y verticales de acuerdo a la orientación del Pozo. Los resultados de los ensayos fueron clasificados según dicha orientación debido a la alta anisotropía que presentan estas rocas tipo shale que tienen la propiedad de haber organizado sus granos en forma lineal hasta formar planos durante su depositación, figura 5-1.



Figura 4-1. Roca shale

Por esta razón la anisotropía se ve reflejada en su resistencia y deformación, siendo necesario realizar análisis por separado a las muestras verticales y muestras horizontales.

El primer paso fue recolectar toda la información, iniciando por los ensayos de laboratorio que en este caso corresponde al módulo de elasticidad, también conocido como el módulo de Young. Adicionalmente se clasificó de acuerdo a su orientación, formación y profundidad. Tabla 4-1. Las profundidades tanto en los datos de laboratorio como en los registros fueron eliminadas para conservar la confidencialidad de los resultados.

**Tabla 4-1** Densidad de datos módulo Elasticidad

<b>N</b>	<b>Orientación</b>	<b>Formación</b>	<b>Miembro</b>	<b>Módulo de Young</b>
1	Vertical	La Luna	Galembo	1.1222E+06
2	Horizontal	La Luna	Galembo	3.3551E+06
3	Horizontal	La Luna	Galembo	9.2290E+05
4	Horizontal	La Luna	Galembo	2.0214E+06
5	Vertical	La Luna	Galembo	2.4339E+06
6	Horizontal	La Luna	Galembo	1.7651E+06
7	Horizontal	La Luna	Galembo	3.5529E+06
8	Horizontal	La Luna	Galembo	3.969E+06
9	Horizontal	La Luna	Galembo	5.248E+06
10	Horizontal	La Luna	Galembo	4.179E+06
11	Horizontal	La Luna	Galembo	1.806E+06
12	Horizontal	La Luna	Galembo	1.6687E+06
13	Horizontal	La Luna	Galembo	5.714E+06
14	Horizontal	La Luna	Galembo	1.5147E+06
15	Horizontal	La Luna	Galembo	1.1336E+06
16	Horizontal	La Luna	Galembo	9.8968E+05
17	Horizontal	La Luna	Galembo	1.9118E+06
18	Horizontal	La Luna	Galembo	3.4630E+06
19	Horizontal	La Luna	Galembo	1.4762E+06
20	Horizontal	La Luna	Galembo	6.243E+06
21	Horizontal	La Luna	Galembo	1.964E+06
22	Horizontal	La Luna	Galembo	3.983E+06
23	Horizontal	La Luna	Galembo	6.365E+06
24	Horizontal	La Luna	Galembo	2.046E+06
25	Horizontal	La Luna	Galembo	3.552E+06
26	Horizontal	La Luna	Galembo	4.039E+06
27	Horizontal	La Luna	Galembo	1.449E+06
28	Horizontal	La Luna	Galembo	3.368E+06
29	Horizontal	La Luna	Galembo	1.307E+06
30	Horizontal	La Luna	Galembo	2.274E+06
31	Horizontal	La Luna	Galembo	1.837E+06
32	Vertical	La Luna	Galembo	5.011E+06
33	Vertical	La Luna	Galembo	3.565E+06
34	Horizontal	La Luna	Galembo	5.746E+06
35	Horizontal	La Luna	Galembo	7.801E+05
36	Horizontal	La Luna	Galembo	4.424E+06
37	Vertical	La Luna	Galembo	2.021E+06
38	Horizontal	La Luna	Galembo	3.643E+06
39	Horizontal	La Luna	Galembo	7.352E+06
40	Vertical	La Luna	Galembo	2.191E+06
41	Vertical	La Luna	Galembo	2.260E+06

<b>N</b>	<b>Orientación</b>	<b>Formación</b>	<b>Miembro</b>	<b>Módulo de Young</b>
42	Horizontal	La Luna	Galemo	3.400E+06
43	Horizontal	La Luna	Pujamana	2.711E+06
44	Horizontal	La Luna	Pujamana	6.598E+06
45	Vertical	La Luna	Pujamana	6.369E+05
46	Horizontal	La Luna	Pujamana	2.819E+06
47	Horizontal	La Luna	Salada	4.839E+06
48	Horizontal	La Luna	Salada	5.767E+06
49	Horizontal	La Luna	Salada	8.648E+06
50	Horizontal	La Luna	Salada	4.882E+06
51	Vertical	La Luna	Salada	6.573E+06
52	Vertical	La Luna	Salada	3.982E+06
53	Horizontal	La Luna	Salada	1.508E+06
54	Horizontal	La Luna	Salada	4.951E+06
55	Vertical	Simití	FALSO	4.898E+06
56	Vertical	Tablazo	FALSO	3.6986E+05
57	Vertical	Tablazo	FALSO	8.0841E+05
58	Vertical	Tablazo	FALSO	1.792E+06
59	Vertical	Tablazo	FALSO	7.3664E+05
60	Vertical	Tablazo	FALSO	1.260E+06
61	Vertical	Tablazo	FALSO	9.1951E+05
62	Vertical	Tablazo	FALSO	8.6551E+05
63	Vertical	Tablazo	FALSO	1.1135E+06
64	Vertical	Tablazo	FALSO	6.8636E+05
65	Vertical	Tablazo	FALSO	1.1904E+06

Luego de recolectar la data anterior que corresponden a ensayos de laboratorio, se reunió la data proveniente de registros de pozo. Los registros utilizados fueron el Gamma Ray, Densidad, Porosidad, Resistivo, tiempo de transito Onda de corte (DT Shear) y tiempo de transito onda compresional (DT Compresional).

**Tabla 4-2.** Datos de Registros formación La Luna

<b>N</b>	<b>Densidad</b>	<b>GR_EDIT</b>	<b>Porosidad</b>	<b>DT Compresional</b>	<b>DT Shear</b>	<b>Resistividad</b>
1	2.3775	78.199	0.368	97.839	187.052	24.939
2	2.6438	56.647	0.190	85.893	164.935	13.437
3	2.6438	56.647	0.190	85.893	164.935	13.437
4	2.4495	65.528	0.222	86.895	157.634	7.469
5	2.4795	75.731	0.239	83.778	149.302	12.115
6	2.4795	75.731	0.239	83.778	149.302	12.115
7	2.4916	92.072	0.260	84.822	158.723	11.026
8	2.4905	69.261	0.215	80.756	144.823	29.112

N	Densidad	GR_EDIT	Porosidad	DT Compresional	DT Shear	Resistividad
9	2.4629	87.830	0.231	84.251	153.190	16.948
10	2.4910	85.699	0.236	87.091	156.477	14.851
11	2.4910	85.699	0.236	87.091	156.477	14.851
12	2.4839	71.792	0.200	86.026	156.831	16.773
13	2.5119	85.711	0.209	83.685	153.067	10.086
14	2.5029	68.956	0.168	83.412	154.518	11.602
15	2.5029	68.956	0.168	83.412	154.518	11.602
16	2.4834	88.185	0.192	90.304	160.968	13.994
17	2.5020	68.438	0.190	85.425	153.332	10.560
18	2.4786	77.800	0.210	90.156	161.769	11.875
19	2.4786	77.800	0.210	90.156	161.769	11.875
20	2.4644	53.102	0.179	86.237	148.404	10.478
21	2.4649	59.189	0.190	93.439	163.745	17.231
22	2.3386	101.746	0.241	95.099	176.835	5.008
23	2.4353	76.980	0.241	104.237	186.886	8.747
24	2.4983	58.286	0.253	88.322	162.796	8.373
25	2.4114	65.944	0.262	88.884	162.780	10.152
26	2.4141	72.425	0.243	84.361	152.316	7.240
27	2.4158	69.061	0.211	92.815	166.270	39.990
28	2.4229	58.477	0.212	87.716	149.039	12.777
29	2.4100	43.164	0.198	85.009	143.906	24.086
30	2.5129	107.001	0.181	71.293	130.379	19.535
31	2.4792	45.417	0.204	84.691	152.534	9.699
32	2.5225	48.346	0.113	85.326	208.673	337.510
33	2.3714	123.594	0.222	81.744	140.091	34.575
34	2.5199	82.801	0.117	66.419	120.921	133.342
35	2.4947	74.840	0.190	78.102	138.659	159.533
36	2.4378	101.582	0.178	80.770	146.749	27.149
37	2.5428	63.508	0.114	85.721	160.162	46.801
38	2.5413	94.204	0.302	79.245	141.611	8.683
39	2.5532	95.660	0.348	89.048	171.531	4.478
40	2.5532	95.660	0.348	89.048	171.531	4.478
41	2.5600	40.642	0.141	98.224	193.371	59.581
42	2.5600	40.642	0.141	98.224	193.371	59.581
43	2.5270	79.846	0.341	96.343	190.775	3.308
44	2.5270	79.846	0.341	96.343	190.775	3.308
45	2.5804	89.172	0.264	90.022	174.560	3.786
46	2.5154	81.368	0.239	84.355	154.673	7.062
47	2.4433	73.253	0.192	85.732	152.392	41.411
48	2.4117	82.583	0.217	87.222	150.515	51.807
49	2.4417	93.825	0.242	84.436	148.243	48.410
50	2.3977	111.086	0.213	86.214	141.457	57.152

N	Densidad	GR_EDIT	Porosidad	DT Compresional	DT Shear	Resistividad
51	2.4793	64.324	0.114	74.510	125.938	59.514
52	2.4960	106.065	0.180	78.273	141.731	50.394
53	2.5136	234.121	0.227	87.524	158.443	21.353
54	2.5285	174.981	0.164	71.728	137.002	424.467
55	2.6264	97.424	0.322	97.062	203.914	1.850
56	2.3425	171.202	-999.250	89.719	177.136	8.141
57	2.2894	181.304	-999.250	106.254	209.998	3.750
58	2.3435	169.440	-999.250	100.691	166.152	3.390
59	2.4359	155.602	-999.250	89.400	151.597	19.412
60	2.3594	138.184	-999.250	94.927	169.567	14.584
61	2.5394	101.110	-999.250	68.911	135.524	139.089
62	2.5711	82.596	-999.250	65.574	130.539	39.530
63	2.4910	66.689	-999.250	65.493	146.894	43.189
64	2.5245	43.531	-999.250	66.199	147.645	7.245
65	2.5245	43.531	-999.250	66.199	147.645	7.245

Luego de obtenidos los registros se procedió a realizar una regresión simple entre el módulo de Young y cada uno de los registros seleccionados, además se determinó el valor del coeficiente de determinación.

### Regresión Simple Muestras Verticales Pozo La luna-1

Se tomó toda la data perteneciente a la orientación vertical del Pozo la luna y se le aplicó regresión lineal simple a cada una de las variables que se extrajeron de los registros, ver anexo 1. Los resultados del coeficiente de determinación fueron:

**Tabla 4-3.** Coeficientes de Determinación Pozo la Luna 1

Variable	Coeficiente de determinación (R <sup>2</sup> )
Módulo de Young Vs Densidad	0.0692
Módulo de Young Vs Gamma Ray	0.1215
Módulo de Young Vs DT corte	8.00E-05
Módulo de Young Vs DT compresional	0.0034
Módulo de Young Vs Resistividad	0.1532

Como los resultados del coeficiente de determinación en la orientación vertical de todo el pozo fueron bajos, en los datos horizontales únicamente se tomaron los clasificados como la formación La Luna y se realizaron las correspondientes gráficas (Anexo 2). La fuerza de correlación se expresa en la Tabla 4-4

**Tabla 4-4.** Coeficientes de Determinación Formación La Luna horizontal

Variable	Coeficiente de Determinación (R <sup>2</sup> )
Módulo de Young Vs Densidad	0.016
Módulo de Young Vs Porosidad	0.0969
Módulo de Young Vs Gamma Ray	0.0144
Módulo de Young Vs DT corte	0.003
Módulo de Young Vs DT compresional	0.0009
Módulo de Young Vs Resistividad	0.0132

De igual manera, los resultados no son satisfactorios, los coeficientes de determinación no son aceptables.

La formación la luna está compuesta por tres miembros que se denominan Galembó, Salada y Pujamana. Teniendo en cuenta que la mayoría de datos pertenecen al miembro Galembó, se grafica el módulo de Young con cada una de las variables provenientes de los registros, pero solamente teniendo en cuenta dicho miembro (anexo 3). El coeficiente de determinación para este caso se muestra en la Tabla 4-5

**Tabla 4-5.** Coeficientes de Determinación miembro Galembó, orientación horizontal.

Variable	Coeficiente de Determinación (R <sup>2</sup> )
Módulo de Young Vs Densidad	0.0028
Módulo de Young Vs Porosidad	0.0839
Módulo de Young Vs Gamma Ray	0.1118
Módulo de Young Vs DT corte	0.0064
Módulo de Young Vs DT compresional	0.0009
Módulo de Young Vs Resistividad	0.0054

Se realizaron las gráficas para el miembro Galembó pero en la orientación Vertical (anexo 4). Los coeficientes de variación obtenidos se observan en la Tabla 4-6

**Tabla 4-6.** Coeficientes de variación miembro Galembó, orientación Vertical

Variable	Coeficiente de Determinación (R <sup>2</sup> )
Módulo de Young Vs Densidad	0.0074
Módulo de Young Vs Gamma Ray	0.0033
Módulo de Young Vs DT corte	0.0421
Módulo de Young Vs DT Compresional	0.3351
Módulo de Young Vs Resistividad	0.6704

En este último análisis si se obtuvieron 2 coeficientes de determinación que mostraron un buen ajuste, estos son el tiempo de transito compresional y la resistividad.

En busca de mejorar el coeficiente de determinación de las variables petrofísicas que se relacionan con el Modulo de Young, se recurre a la Técnica de Regresión múltiple.

#### 4.1 Regresión Múltiple Pozo La luna

La regresión múltiple se aplicó al Pozo la luna (formación La luna, tablazo y Simití), luego se aplicó a solamente a la formación la Luna (miembro Galemo, Pujamana y Salada), finalmente solo se aplica al miembro Galemo.

El análisis Multivariado al pozo la Luna y la Formación la luna solamente se aplicó a una orientación, porque es un área muy grande y se requería comprobar que no se obtenía resultados aceptables, así que no era válido hacer un estudio detallado.

Se aplicó regresión lineal múltiple al pozo la luna-1 en dirección vertical obteniendo los siguientes resultados de la Tabla 4-7.

**Tabla 4-7.** Estadística de la Regresión pozo la luna, orientación vertical

Estadísticas de la Regresión	
Coeficiente de correlación múltiple	0.554484659
Coeficiente de determinación R <sup>2</sup>	0.307453238
R <sup>2</sup> ajustado	0.060115108
Error típico	1681241.548
Observaciones	20

**Tabla 4-8.** Análisis de varianza pozo la luna, orientación vertical

	Coeficientes	Error típico	Estadístico t	Probabilidad	Inferior 95%	Superior 95%
<b>Intercepción</b>	-6504751.508	16282406.496	-0.399	0.696	-41427040.21	28417537.193
<b>Densidad</b>	2941439.312	5961218.132	0.493	0.629	-9844101.982	15726980.606
<b>Gamma Ray</b>	-12445.656	14088.236	-0.883	0.392	-42661.917	17770.606
<b>DT compresional</b>	101288.769	69604.083	1.455	0.168	-47997.142	250574.680
<b>DT corte</b>	-39571.527	29932.414	-1.322	0.207	-103770.170	24627.116
<b>Resistividad</b>	11220.234	6310.230	1.778	0.097	-2313.863	24754.332

Se observa notoriamente una mejoría si se compara el coeficiente de determinación de la regresión múltiple, con la regresión simple de cada una de las variables en forma independiente.

✓ **Regresión Múltiple Formación La luna**

Se aplicó regresión lineal múltiple a la formación la luna teniendo en cuenta los tres miembros (Galembopujamana y salada) en la dirección Horizontal. Los resultados se muestran en la Tabla 4-9

**Tabla 4-9.** Estadística de la Regresión Formación la luna, orientación Horizontal

<b>Estadísticas de la regresión</b>	
Coeficiente de correlación múltiple	0.434794472
Coeficiente de determinación R <sup>2</sup>	0.189046233
R <sup>2</sup> ajustado	0.041600093
Error típico	1946975.3
Observaciones	40

**Tabla 4-10.** Análisis de varianza Formación la luna, orientación Horizontal

	<b>Coeficientes</b>	<b>Error típico</b>	<b>Estadístico t</b>	<b>Probabilidad</b>	<b>Inferior 95%</b>	<b>Superior 95%</b>
<b>Intercepción</b>	21225265.13	22025107.254	0.964	0.342	-23585152.51	66035682.763
<b>Densidad</b>	-8191461.579	7828108.050	-1.046	0.303	-24117867.16	7734944.000
<b>Gamma Ray</b>	-1629.420588	10647.137	-0.153	0.879	-23291.18425	20032.343
<b>Porosidad</b>	17993139.94	8276857.102	2.174	0.037	1153747.548	34832532.324
<b>DT Compresional</b>	-72230.20849	158985.805	-0.454	0.653	-395689.2598	251228.843
<b>DT Corte</b>	30730.08897	71395.476	0.430	0.670	-114525.0998	175985.278
<b>Resistividad</b>	7701.642665	5866.173	1.313	0.198	-4233.176638	19636.462

Si se observa el coeficiente de determinación, los resultados específicamente de esta regresión no son satisfactorios.

✓ **Regresión Múltiple Formación la Luna, miembro Galembopujamana, Orientación Horizontal**

Se realizó regresión Múltiple a los datos correspondientes al miembro Galembopujamana de la formación la luna en la orientación horizontal y se obtuvieron los siguientes resultados:

**Tabla 4-11.** Estadística de la Regresión Formación la luna, Miembro Galembo, orientación Horizontal

<b>Estadísticas de la regresión</b>	
Coeficiente de correlación múltiple	0.404628865
Coeficiente de determinación R <sup>2</sup>	0.163724519
R <sup>2</sup> ajustado	-0.036981597
Error típico	1827758.229
Observaciones	32

**Tabla 4-12.** Análisis de Varianza Formación la luna, Miembro Galembo, orientación Horizontal

	<b>Coeficientes</b>	<b>Error típico</b>	<b>Estadístico t</b>	<b>Probabilidad</b>	<b>Inferior 95%</b>	<b>Superior 95%</b>
<b>Intercepción</b>	3275526.931	23371908.645	0.140	0.890	-44859819.975	51410873.837
<b>Densidad</b>	-2134215.733	7939769.985	-0.269	0.790	-18486478.117	14218046.650
<b>Gamma Ray</b>	27942.688	23570.029	1.186	0.247	-20600.695	76486.071
<b>Porosidad</b>	8331072.277	9277639.965	0.898	0.378	-10776584.909	27438729.463
<b>DT Compresional</b>	-51125.058	186441.574	-0.274	0.786	-435108.668	332858.552
<b>DT Corte</b>	36323.643	83002.292	0.438	0.665	-134622.776	207270.063
<b>resistividad</b>	3493.623	11832.322	0.295	0.770	-20875.501	27862.746

De igual forma los resultados obtenidos no son los mejores.

✓ **Regresión Múltiple Formación la Luna, miembro Galembo, Orientación Vertical**

Se aplicó regresión Múltiple a los datos correspondientes al miembro Galembo de la formación La Luna en la orientación vertical y se obtuvieron los siguientes resultados:

**Tabla 4-13.** Estadística de la Regresión Formación la luna, Miembro Galembo, orientación Vertical.

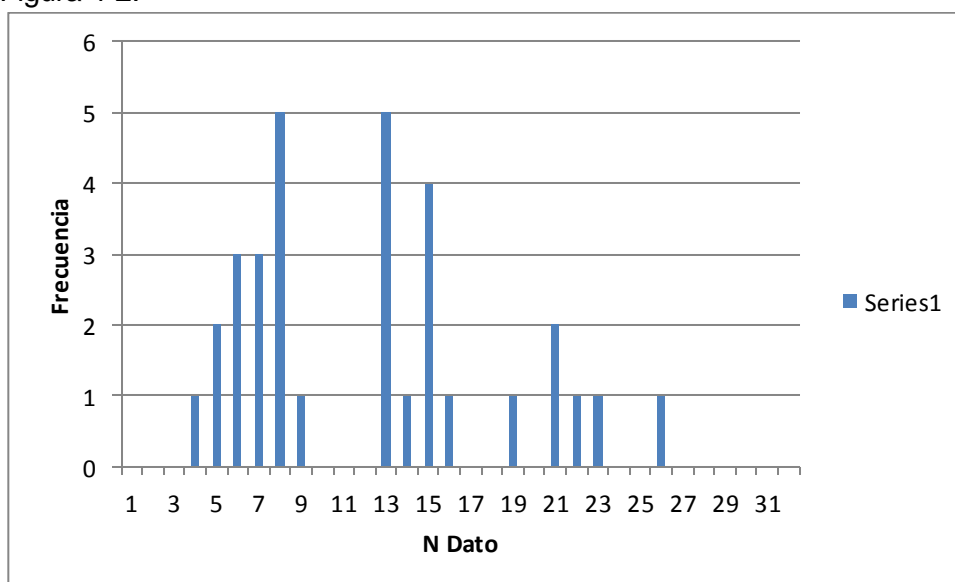
<b>Estadísticas de la regresión</b>	
Coeficiente de correlación múltiple	0.974372362
Coeficiente de determinación R <sup>2</sup>	0.949401499
R <sup>2</sup> ajustado	0.696408996
Error típico	695245.7687
Observaciones	7

**Tabla 4-14.** Análisis de varianza Formación la luna, Miembro Galembó, orientación Vertical

	Coeficientes	Error típico	Estadístico t	Probabilidad	Inferior 95%	Superior 95%
<b>Intercepción</b>	-24555748.090	29342333.37	-0.837	0.556	-397385443.273	348273947.093
<b>Densidad</b>	7917176.172	7590747.153	1.043	0.487	-88532411.253	104366763.597
<b>Gamma Ray</b>	28590.121	23914.225	1.196	0.443	-275268.915	332449.158
<b>DT compresional</b>	212804.237	261347.699	0.814	0.565	-3107933.128	3533541.603
<b>DT corte</b>	-89206.975	85778.862	-1.040	0.488	-1179130.764	1000716.813
<b>Resistividad</b>	25602.059	15599.244	1.641	0.348	-172605.129	223809.247

Luego de realizar regresión múltiple en las dos orientaciones se obtuvieron buenos resultados en la vertical, con los datos que pertenecen solamente a una formación (Galembó). Con el objeto de mejorar la correlación de las variables en la orientación horizontal se decidió hacer un filtro a los datos observados.

Se construyó un gráfico de frecuencia para identificar la distribución de los datos en la Figura 4-2.



**Figura 4-2.** Gráfico de frecuencia del Módulo de Elasticidad-formación la luna, miembro Galembó

Los datos corresponden a una distribución normal con sesgo a la derecha. Conociendo esta información se recurre al uso de la técnica Test de Grubbs para mejorar el ajuste.

#### 4.2 Test de Grubbs formación la luna miembro Galembó, Orientación horizontal.

Con el objeto de realizar un filtro a los datos, para mejorar la correlación se aplicó el tratamiento estadístico conocido como test de Grubbs a los datos de la formación la luna, miembro Galembó en la orientación horizontal. Se Siguió la

metodología empleada con los datos de UCS Y Modulo de Young del artículo estudiado en el capítulo 3. En la orientación vertical del miembro galemba no se aplicó el test de Grubbs debido a que su coeficiente de determinación dió igual a 0.949, siendo un buen ajuste.

Se determinó el T tabulado con 95% de probabilidad y 5% de riesgo para un n igual a 32 que corresponde a la data del miembro Galemba de la formación La Luna. T=1.6966 Los T calculados se presentan en la siguiente tabla

**Tabla 4-15.** T calculados data Pozo la Luna

<b>N</b>	<b>Módulo de Young</b>	<b>Densidad</b>	<b>GR_EDIT</b>	<b>Porosidad</b>	<b>DT Compresional</b>	<b>DT Shear</b>	<b>Resistivo</b>
1	1.2646	2.9827	1.0155	0.5576	0.0355	0.6121	0.2999
2	0.6526	0.5322	0.4995	0.1745	0.1068	0.1047	0.4745
3	0.7954	0.0105	0.0934	0.5584	0.3358	0.4743	0.3385
4	0.2006	0.2294	1.0429	1.0476	0.1876	0.1804	0.3704
5	0.4326	0.2095	0.2826	0.0215	0.7650	0.7856	0.1589
6	1.1453	0.2891	0.7964	0.3663	0.2686	0.2041	0.1971
7	0.5496	0.2190	0.6726	0.5032	0.1348	0.0243	0.2585
8	0.8491	0.0899	0.1355	0.3209	0.0166	0.0489	0.2022
9	1.4044	0.5964	0.6732	0.1349	0.3490	0.2127	0.3979
10	0.9349	0.4335	0.3003	1.0622	0.3878	0.1118	0.3535
11	1.2274	0.0809	0.8170	0.5109	0.5911	0.3364	0.2835
12	0.7137	0.4183	0.3304	0.5578	0.1019	0.1942	0.3841
13	0.1506	0.0046	0.2136	0.0978	0.5700	0.3921	0.3456
14	0.9564	0.0046	0.2136	0.0978	0.5700	0.3921	0.3456
15	1.6993	0.2615	1.2215	0.7996	0.0133	0.5367	0.3865
16	0.6847	0.2526	0.8678	0.5637	1.0363	0.5294	0.1888
17	0.4403	2.5379	1.6050	0.6080	1.2721	1.4391	0.5465
18	1.7672	0.7893	0.1659	0.6038	2.5700	2.1376	0.4371
19	0.6391	0.3506	0.9203	0.8881	0.3096	0.4635	0.4481
20	0.2001	1.2218	0.4753	1.0800	0.3894	0.4624	0.3960
21	0.4716	1.1717	0.0987	0.6544	0.2530	0.2648	0.4812
22	0.9716	1.1405	0.2942	0.0832	0.9478	0.7048	0.4773
23	0.0979	1.0124	0.9092	0.0580	0.2235	0.4926	0.3192
24	1.0506	1.2460	1.7990	0.3861	0.1610	0.8493	0.0118
25	0.5117	0.6158	1.9104	0.7742	2.1091	1.7893	0.1214
26	0.7552	0.0060	1.6681	0.2452	0.2062	0.2497	0.4092
27	1.4223	0.7413	0.5042	2.2359	2.8013	2.4467	3.2096
28	1.3442	0.2855	0.0416	0.5509	1.1420	1.2140	3.9762
29	0.6858	0.7443	1.5955	0.8288	0.7630	0.6517	0.1015

N	Módulo de Young	Densidad	GR_EDIT	Porosidad	DT Compresional	DT Shear	Resistivo
30	0.2509	1.1282	1.1668	2.0049	0.9796	1.0088	0.4390
31	2.3174	1.3433	1.2513	3.0439	0.4126	1.0705	0.5621
32	0.1156	1.4678	1.9455	1.6888	1.7159	2.5882	1.0508

Los datos para la regresión múltiple que corresponden al 95% de la probabilidad en el test de Grubbs se redujeron de 32 a 21 observaciones. Los resultados estadísticos con esta condición se observan en la Tabla 4-16, y los coeficientes y las variables involucradas están en la Tabla 4-17. El coeficiente de determinación mejoró paso de 0.163 (Tabla 4-11) a 0.49.

**Tabla 4-16.** Estadística de la Regresión con 95% de probabilidad Test de Grubbs

Estadísticas de la regresión	
Coefficiente de correlación múltiple	0.700636124
Coefficiente de determinación R <sup>2</sup>	0.490890979
R <sup>2</sup> ajustado	0.272701398
Error típico	1178317.075
Observaciones	21

**Tabla 4-17.** Análisis de variables con 95% de probabilidad Test de Grubbs

	Coefficientes	Error típico	Estadístico t	Probabilidad	Inferior 95%	Superior 95%
<b>Intercepción</b>	24586765.89	27511756.4	0.894	0.387	-34420083	83593614.8
<b>Densidad</b>	-5187597.146	10017945.2	-0.518	0.613	-26673952.6	16298758.29
<b>Gamma Ray</b>	42532.84453	21984.5802	1.935	0.073	-4619.39037	89685.07944
<b>Porosidad</b>	18022873.06	11899173.2	1.515	0.152	-7498315.12	43544061.24
<b>DT Compresional</b>	-16677.4026	184038.428	-0.091	0.929	-411400.573	378045.7678
<b>DT Shear</b>	-92641.17214	104018.547	-0.891	0.388	-315738.766	130456.4221
<b>Resistivo</b>	-1953.211647	37613.3773	-0.052	0.959	-82625.8827	78719.45937

Si se aplica test de Grubbs con un riesgo del 5% para las variables independientes y 10% para la variable dependiente, el número de observaciones se convierte en 20 y se obtienen los siguientes resultados en la regresión múltiple:

**Tabla 4-18.** Estadística de la Regresión con 95% de probabilidad variables independientes y 90 % variable dependiente en Test de Grubbs

<b>Estadísticas de la regresión</b>	
Coeficiente de correlación múltiple	0.756364492
Coeficiente de determinación R <sup>2</sup>	0.572087245
R <sup>2</sup> ajustado	0.37458905
Error típico	988033.1197
Observaciones	20

El coeficiente de determinación mejoró, paso de 0.49 a 0.572. Se obtuvo una mejora con respecto al análisis anterior del 14%, con solamente la reducción del número de observaciones de 21 a 20.

**Tabla 4-19.** Análisis de variables con 95% de probabilidad variables independientes y 90 % variable dependiente en Test de Grubbs

	<b>Coeficientes</b>	<b>Error típico</b>	<b>Estadístico t</b>	<b>Probabilidad</b>	<b>Inferior 95%</b>	<b>Superior 95%</b>
<b>Intercepción</b>	35125455.5	23414629.8	1.500	0.157	-15458776.83	85709687.7
<b>Densidad</b>	-9863147.35	8586365.26	-1.149	0.271	-28412861.73	8686567.03
<b>Gamma Ray</b>	33841.0038	18728.4583	1.807	0.094	-6619.370391	74301.378
<b>Porosidad</b>	19018189.7	9984780.58	1.905	0.079	-2552617.341	40588996.7
<b>DT Compresional</b>	-8885.58687	154346.906	-0.058	0.955	-342331.8042	324560.63
<b>DT Shear</b>	-89532.5567	87228.8241	-1.026	0.323	-277978.9741	98913.8608
<b>Resistivo</b>	7069.27109	31725.4389	0.223	0.827	-61469.37266	75607.9149

### 4.3 Identificación de variables estadísticamente significativas:

En este punto se identificaron aquellas variables independientes que tenían relación con la variable dependiente para construir un modelo simplificado. Por tanto se buscó conocer que contribución tenía cada variable independiente en la predicción de la variable dependiente.

El procedimiento que se realiza es la eliminación de las variables independientes que no son estadísticamente significativas utilizando la columna de la probabilidad. En las tablas que se han venido trabajando, se elimina el valor P más alto. Y en algunas estadísticas utilizan el valor de 1-p, en este caso las variables que no son significativas son las del valor más bajo.

Si se observa la Tabla 4-19 los valores más altos en la columna de probabilidad los tiene el DT compresional y la Resistividad, 0.955 y 0.827, respectivamente.

Los resultados que se obtuvieron eliminando estas dos variables no cambian, observar Tabla 4-20.

**Tabla 4-20.** Estadística de la Regresión eliminando dos variables no significativas

<b>Estadísticas de la regresión</b>	
Coefficiente de correlación múltiple	0.75522948
Coefficiente de determinación R <sup>2</sup>	0.570371567
R <sup>2</sup> ajustado	0.455803985
Error típico	921650.8777
Observaciones	20

**Tabla 4-21.** Análisis eliminando dos variables no significativas.

	<i>Coefficientes</i>	<i>Error típico</i>	<i>Estadístico t</i>	<i>Probabilidad</i>	<i>Inferior 95%</i>	<i>Superior 95%</i>
<b>Intercepción</b>	36218432.3	18575763	1.950	0.070	-3374869.233	75811733.8
<b>Densidad</b>	-10295632	7026544.86	-1.465	0.163	-25272357.86	4681093.84
<b>Porosidad</b>	35254.3857	15962.1742	2.209	0.043	1231.816821	69276.9546
<b>GR</b>	18420120.4	8189523.79	2.249	0.040	964563.596	35875677.1
<b>DT Shear</b>	-93780.6777	35488.0957	-2.643	0.018	-169421.7632	-18139.5922

Si se continúa eliminando la variable que sigue en cuanto a la probabilidad más alta, que en este caso corresponde a la densidad (0.163), el coeficiente de determinación se reduce pasando de 0.57 a 0.50.

**Tabla 4-22.** Estadística de la Regresión eliminando tres variables no significativas

<b>Estadísticas de la regresión</b>	
Coefficiente de correlación múltiple	0.713357404
Coefficiente de determinación R <sup>2</sup>	0.508878786
R <sup>2</sup> ajustado	0.416793559
Error típico	954113.2309
Observaciones	20

**Tabla 4-23.** Análisis eliminando tres variables no significativas

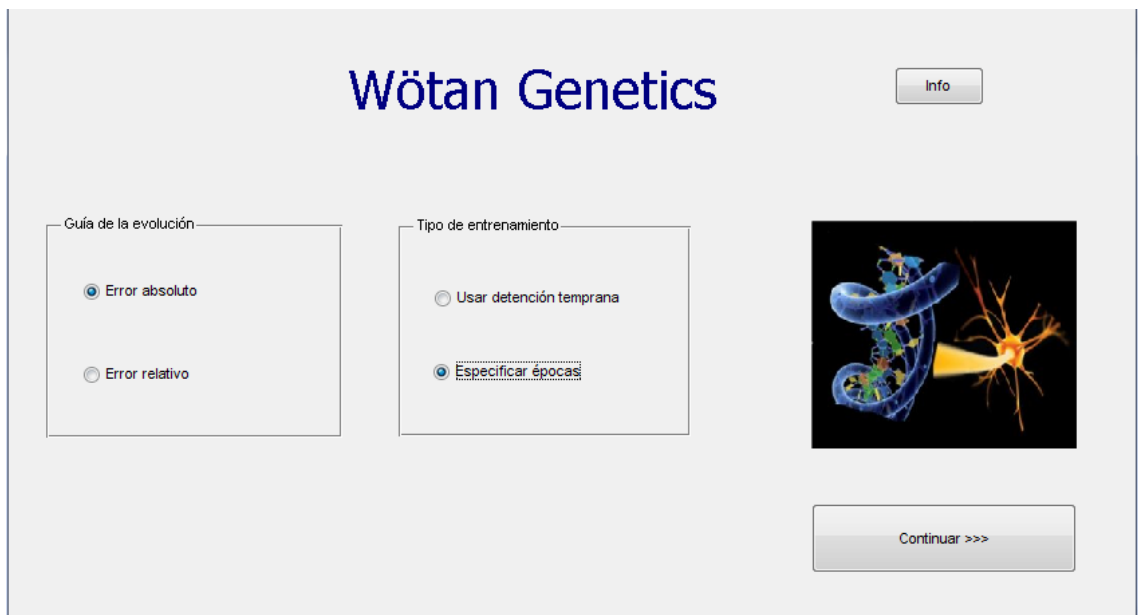
	<i>Coefficientes</i>	<i>Error típico</i>	<i>Estadístico t</i>	<i>Probabilidad</i>	<i>Inferior 95%</i>	<i>Superior 95%</i>
<b>Intercepción</b>	10348548.6	5977239.25	1.731	0.102	-2322632.581	23019729.7
<b>Porosidad</b>	33766.9795	16490.9439	2.047	0.057	-1192.259906	68726.2188
<b>GR</b>	20755454.6	8315871.36	2.495	0.023	3126594.868	38384314.4
<b>DT Shear</b>	-93304.1907	36736.5137	-2.539	0.021	-171182.1207	-15426.2606

En general con el análisis multivariado se obtuvieron mejores resultados que con una sola variable; y además aplicando el Test de Grubbs para eliminar datos anómalos mejora notablemente la predicción.

## 5 APLICACIÓN DE LAS REDES NEURONALES A SHALES PARA EL CÁLCULO DEL MÓDULO DE YOUNG (E)

La Herramienta utilizada para la predicción de variables a través de redes Neuronales fue un Sistema Automático de entrenamiento de redes neuronales Artificiales basado en ajuste Genético de Parámetros y variación de arquitectura, desarrollado en MATLAB 7.5.0 R2007b. El nombre del software es ORION y es propiedad Ecopetrol, creada en el instituto colombiano del petróleo.

La interfaz de usuario principal del ORION, Figura 5-1, muestra dos maneras de realizar el entrenamiento, una de ellas es detención temprana y la otra especificación de Épocas.



**Figura 5-1.** Interface de usuario principal ORION  
**Fuente:** Herramienta ORION

El número de épocas es el número de veces en las que le serán mostrados los datos de entrenamiento a la red.

Al escoger este tipo de entrenamiento se abre una interfaz (Figura 5-2) y solicita dos grupos de datos, entrenamiento y prueba cada una con sus respectivas entradas y salidas. Adicionalmente se solicita el número de épocas para los entrenamientos.



**Figura 5-2.** Interface de carga de Datos  
**Fuente:** Herramienta ORION

Cuando se usa detección temprana se agrega un grupo de datos adicional en el entrenamiento, estos datos se utilizan para detener el entrenamiento en el instante en el que las respuestas de la red se alejen demasiado de las respuestas de los datos de verificación.

La interfaz permite escoger el tipo de normalización que se desea aplicar a los datos de entrada y salida de los entrenamientos. Las Opciones de normalización que tiene el programa son:

- ✓ Ninguna: No realiza ninguna acción sobre los datos.
- ✓ Sigmoidal: Este tipo de normalización ajusta los datos en un rango entre cero y uno  $[0,1]$ .
- ✓ Tangente Hiperbólica: Este tipo de normalización ajusta los datos en un rango entre menos uno y uno  $[-1,1]$ .

Para las predicciones de este trabajo se escogió el tipo de entrenamiento Especificación por épocas, y el tipo de normalización Tangente Hiperbólica.

Los datos que se utilizaron para la predicción por redes neuronales hacen parte del miembro Galemo de la formación La Luna en la orientación Horizontal.

El primer paso fue encontrar aquellos valores mínimos y valores máximos para tenerlos en cuenta en el grupo de datos de entrenamiento, esto debido a que se ha comprobado que las redes neuronales son buenas interpolando y malas extrapolando.

Toda la data de las corridas se organizó en 2 grupos, estos fueron: un conjunto de 25 para entrenamiento y otro de 7 datos para prueba que corresponde

aproximadamente al 80% y 20%, respectivamente. Los resultados de los Coeficientes de variación de algunas de las corridas realizadas fueron:

**Tabla 5-1.** Coeficientes de Determinación diferentes redes neuronales

RED	Neuronas y Capas Ocultas	R <sup>2</sup>
1	6	0.077
2	2 8 8	0.0096
3	4 1 9 9	0.244
4	2 7 4 9	0.144
5	7 9 9	0.405
6	3 3 9	0.501
7	3	0.45
8	6 4 1 6	0.488
9	3 1	0.641

La tabla anterior además da información de la estructura de las redes neuronales y capas ocultas, por ejemplo en la fila 2 significa que tiene tres capas ocultas, en la primera hay una configuración de 2 neuronas, en la segunda capa 8 neuronas y en la última capa de nuevo otras 8.

La red número 9 de la Tabla 5-1 obtuvo un coeficiente de determinación alto si se compara con los datos de regresión simple, incluso los de regresión múltiple. Esto se logró con una selección de datos específicos de entrenamiento, es decir los datos de entrenamiento y prueba fueron variando en cada corrida buscando obtener un mejor ajuste. En efecto los datos de entrenamiento y prueba para la red 9 son los mostrados en la Tabla 5-2. Datos de entrenamiento y prueba con coeficiente de determinación más alto. Tabla 5-2

**Tabla 5-2.** Datos de entrenamiento y prueba con coeficiente de determinación más alto.

Densidad (g/cm <sup>3</sup> )	GR_EDIT (GAPI)	Porosidad (Frac)	DT Compresional (μs/F)	DTShear (μs/F)	Resistivo (HOMM)	Módulo de Young (PSI)
<b>Datos entrenamiento</b>						
2.541	94.204	0.302	79.245	141.611	8.683	3.643E+06
2.464	53.102	0.179	86.237	148.404	10.478	6.243E+06
2.644	56.647	0.190	85.893	164.935	13.437	9.2290E+05
2.339	101.746	0.241	95.099	176.835	5.008	3.983E+06
2.435	76.980	0.241	104.237	186.886	8.747	6.365E+06
2.513	107.001	0.181	71.293	130.379	19.535	2.274E+06
2.520	82.801	0.117	66.419	120.921	133.342	5.746E+06
2.495	74.840	0.190	78.102	138.659	159.533	7.801E+05

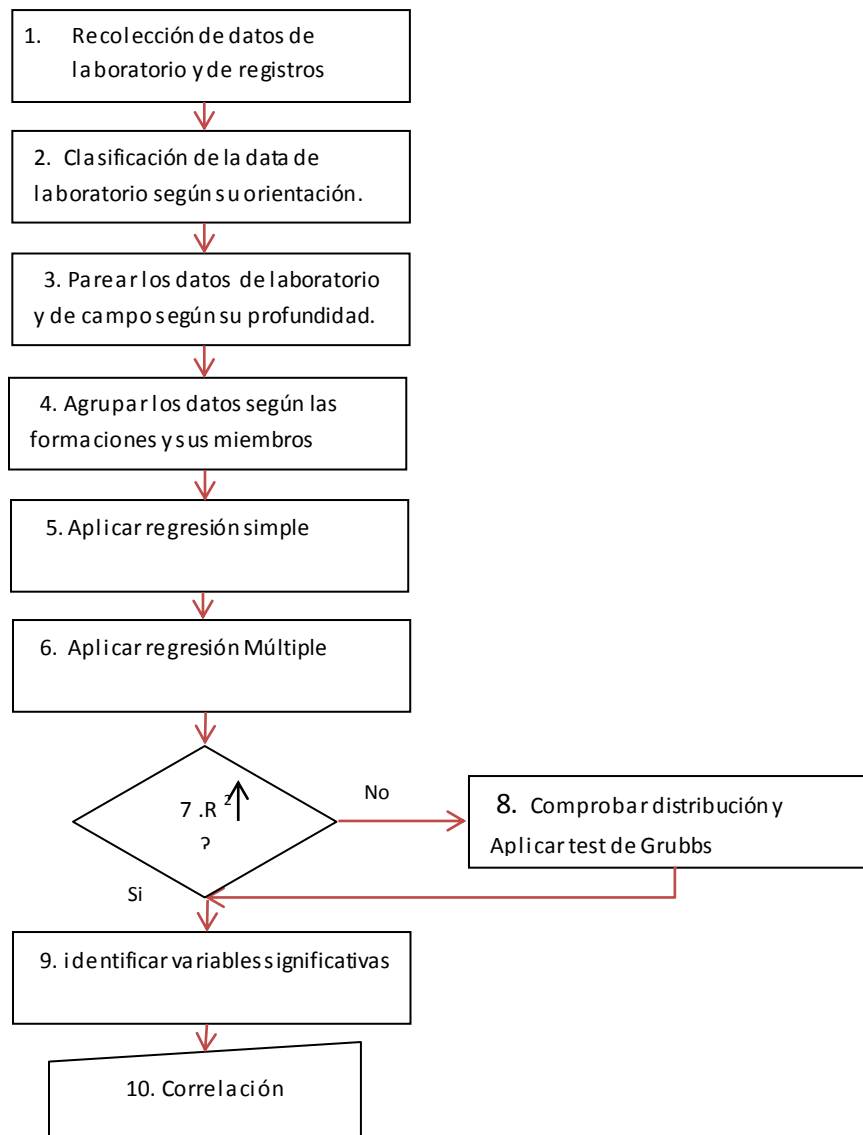
Densidad (g/cm3)	GR_EDIT (GAPI)	Porosidad (Frac)	DT Compresional (μs/F)	DTShear (μs/F)	Resistivo (HOMM)	Módulo de Young (PSI)
2.553	95.660	0.348	89.048	171.531	4.478	7.352E+06
2.560	40.642	0.141	98.224	193.371	59.581	3.400E+06
2.449	65.528	0.222	86.895	157.634	7.469	2.0214E+06
2.479	75.731	0.239	83.778	149.302	12.115	1.7651E+06
2.490	69.261	0.215	80.756	144.823	29.112	3.969E+06
2.463	87.830	0.231	84.251	153.190	16.948	5.248E+06
2.484	71.792	0.200	86.026	156.831	16.773	1.6687E+06
2.512	85.711	0.209	83.685	153.067	10.086	5.714E+06
2.479	77.800	0.210	90.156	161.769	11.875	1.4762E+06
2.465	59.189	0.190	93.439	163.745	17.231	1.964E+06
2.498	58.286	0.253	88.322	162.796	8.373	2.046E+06
2.411	65.944	0.262	88.884	162.780	10.152	3.552E+06
2.416	69.061	0.211	92.815	166.270	39.990	1.449E+06
2.423	58.477	0.212	87.716	149.039	12.777	3.368E+06
2.410	43.164	0.198	85.009	143.906	24.086	1.307E+06
2.479	77.800	0.210	90.156	161.769	11.875	3.4630E+06
2.438	101.582	0.178	80.770	146.749	27.149	4.424E+06
<b>Datos Prueba</b>						
2.502	68.438	0.190	85.425	153.332	10.560	1.9118E+06
2.492	92.072	0.260	84.822	158.723	11.026	3.5529E+06
2.491	85.699	0.236	87.091	156.477	14.851	4.179E+06
2.483	88.185	0.192	90.304	160.968	13.994	9.8968E+05
2.438	101.582	0.178	80.770	146.749	27.149	4.424E+06
2.503	68.956	0.168	83.412	154.518	11.602	1.5147E+06
2.479	45.417	0.204	84.691	152.534	9.699	1.837E+06

## 6 ANALISIS DE RESULTADOS

### 6.1 RESULTADOS ANALISIS MULTIVARIADO

El resultado que se obtuvo de la regresión multivariada es una metodología, que se fue construyendo con los datos del artículo estudiado en el capítulo 3 y los datos del pozo La Luna que se estudió en el capítulo 4.

#### 6.1.1 Metodología para determinar Parámetros geomecánicos a Partir de Registros petrofísicos y pruebas de laboratorio.



1. **Recolección de datos de laboratorio y de registro:** los registros de pozo a utilizar en el análisis no deben provenir de cálculos, sino deben ser el valor original registrado por la herramienta, de esta manera se garantiza menos ruido introducido al modelo.
2. **clasificación de la data de laboratorio según la Orientación:** la anisotropía en los parámetros elásticos de las rocas tipo shale sugieren tratamiento estadísticos por separado. El ajuste de la regresión Múltiple en los datos verticales mostró mejores resultados, este debe ser por la forma en que se toma los registros de pozo y que se realizan a lo largo del eje vertical.
3. **Parear los datos de laboratorio y de campo según su profundidad:** la mayoría de registros tienen datos cada 0.25 y 0.5 pies, así que el número de observaciones están condicionadas es por la cantidad de datos de laboratorio.
4. **Agrupar los datos según las formaciones y los miembros:** las correlaciones siempre mostraron un mejor ajuste en la medida que se estudiaban los datos de una unidad litoestratigráfica menor. Inicialmente se realizó a todo el pozo, luego la formación La Luna y finalmente en el miembro Galembó se obtuvieron mejores ajustes.
5. **Aplicar regresión simple:** Con la regresión simple se tiene una idea de que tan desviada o acertada está cada variable independiente al tratar de predecir, o quizás en este paso pueda descartarse la regresión múltiple si la variable a predecir es descrita por otra únicamente.
6. **Aplicar regresión Múltiple:** la regresión múltiple demostró que efectivamente proporciona correlaciones con mejores ajustes. Esto comprueba que las variables están afectadas por varios parámetros lo cual es más cercano de la realidad.
7. **¿Tiene la regresión un coeficiente de determinación ( $R^2$ ) aceptable? ¿Es mayor a 0.8 por lo menos?** En el caso de los datos del miembro Galembó en la orientación vertical se obtuvo un valor de  $R^2$  igual a 0.949. En este caso no se realizó ningún filtro a los datos, aunque se puede pensar que siempre se debe hacer algún control de calidad, solo que con un  $R^2$  alto seguramente todos los valores están adentro en un riesgo del 10%.
8. **Comprobar la distribución y realizar test de Grubbs:** En este paso se comprueba el tipo de distribución con un diagrama de frecuencia. La mayoría de fenómenos de la naturaleza tienen un comportamiento de distribución normal, a la cual es válido aplicar t –student, que en efecto es en lo que se basa el Test de

Grubbs. Este test es un filtro que se realiza a los datos para mejorar el ajuste del modelo. Inicialmente se debe probar con un riesgo del 5% aplicando test de Grubbs a todas las variables, tanto independientes como la variable a predecir. Si no funciona o no es notorio el cambio en el ajuste, se debe aplicar 5% de riesgo a las variables independientes y 10% en la variable dependiente. Por supuesto, en la medida que aumentamos el riesgo tendería a mejorar el ajuste, pero de la misma manera se van reduciendo la cantidad de datos.

9. **Identificar variables significativas:** se identifican aquellas variables que realmente no están aportando en la descripción del modelo a través de la correlación. En este paso se busca establecer un modelo simplificado.

10. **Correlación:** Finalmente se obtiene una correlación con múltiples variables, pero simplificada que explica la predicción la variable de interés.

Con la regresión múltiple se obtuvieron coeficientes de variación ( $R^2$ ), de 0.572 y 0.949 en el Análisis realizado a los datos de la luna, luego de haber obtenido,  $R^2$ , en la regresión simple por el orden de 0.001 e incluso valores menores.

Con los datos del artículo “Predicción de la resistencia compresiva uniaxial y el módulo de Elasticidad para muestras de Travertino Usando Regresión y Redes Neuronales artificiales” se validaron los resultados y aplicó test de Grubbs filtrando datos no significativos. El  $R^2$  pasó de 0.64 a 0.72, mejorando en 11% el ajuste en la predicción de la resistencia compresiva, y en cuanto al Módulo de Young el  $R^2$  incremento de 0.56 a 0.709, mejorando en un 21%.

### 6.1.2 Ventajas del Análisis Multivariado

Estas son algunas ventajas que se detectaron con el desarrollo del trabajo.

- ✓ Obtienen buenos ajustes con pocos datos.
- ✓ Se puede tener más control sobre las variables, de tal forma que se identifican las que verdaderamente están aportando y mejoran el ajuste de la correlación y fácilmente se eliminan aquellas que no son significativas estadísticamente.
- ✓ Presentan una correlación con las que se puede predecir datos incluso sin programas de cálculo robustos.
- ✓ El resultado se obtiene al instante a comparación de las redes, que puede durar horas, e incluso días en su entrenamiento, además se debe probar con

varios set de datos de entrenamiento para determinar la red que predice con exactitud. Finalmente Obtener una buena red puede tardar quizás 1 semana.

## 6.2 RESULTADOS CON LAS REDES NEURONALES

Las redes neuronales tienen la ventaja que pueden obtener predicciones con coeficientes de variación altos. Si se comparan los  $R^2$  durante el desarrollo de este estudio, siempre que se obtuvieron por los dos métodos; los calculados a partir de las redes fueron más altos que los calculados por la regresión múltiple.

**Tabla 6-1.** Comparación coeficientes de Variación entre regresión Múltiple y Redes Neuronales

Datos	$R^2$ (Regresión Múltiple)		$R^2$ (Redes Neuronales)			
	UCS	Young	UCS		Young	
originales articulo	0.64	0.56	0.86	0.93	0.77	0.82
Calculados articulo	0.64	0.56	0.88		0.76	
calculados articulo datos filtrados	0.73	0.71	0.88		0.76	
Galembo Horizontal	-	0.57	-		0.64	

En la Tabla 6-1 se observa el cuadro comparativo de los resultados de los coeficientes de variación entre la regresión múltiple y las redes neuronales.

Siempre se obtuvieron más altos coeficientes de variación en las redes, aunque si comparamos los datos del miembro Galembo se obtuvo un  $R^2$  de 0.57 y 0.64, con el análisis multivariado y con las redes, respectivamente; esto es una diferencia del 11%. Habría que decidir cuál técnica utilizar en el caso práctico si se necesita con bastante urgencia un resultado, porque el entrenamiento para encontrar una red con buen ajuste es demorado. Mientras con el multivariado aplicando test de Grubbs se puede obtener en aproximadamente 15 minutos.

Las redes tienen un problema, que con pocos datos no se puede aplicar, por ejemplo con la formación Galembo en la orientación vertical, el resultado del  $R^2$  fue 0.95, pero solamente habían 7 datos disponibles de pruebas de laboratorio. Con las redes si se decide aplicar el 80% de entrenamiento y el 20% de prueba. Serian 6 datos de entrenamiento y 1 de prueba, lo que hace imposible determinar un  $R^2$  con un solo dato.

## 7 CONCLUSIONES

- ✓ La aplicación de la técnica de Análisis multivariado demostró que podía explicar mejor el comportamiento de la variable dependiente a la hora de realizar una predicción, que cuando se usa la regresión simple, demostrando que en la naturaleza no ocurren fenómenos aislados, sino por el contrario siempre están afectados por más de un parámetro.
- ✓ La validación de los resultados de la regresión múltiple se realizó mediante datos reportados en la literatura a través de un análisis de varianza de dos muestras por el cual se logró explicar que los resultados eran similares.
- ✓ Los resultados de los análisis de ensayos en laboratorio muchas veces no son representativos de una roca intacta sino por el contrario se ven afectados por alguna fractura, oquedad o un plano que hacen que dichos datos se encuentren dispersos, complicando el ajuste de los mismos; en este caso es factible la aplicación de un método estadístico que permita hacer un filtro para eliminar valores atípicos y mejorar el modelo en términos de la varianza de sus errores. Aplicando test de Grubbs al trabajo de DEHGHAN S et al (2010), el  $R^2$  pasó de 0.64 a 0.72, mejorando en 11% el ajuste en la predicción de la resistencia compresiva, y en cuanto al Módulo de Young el  $R^2$  incremento de 0.56 a 0.709, mejorando en un 21%.
- ✓ Las redes neuronales demostraron ser buenas en la predicción de variables obteniendo coeficientes de variación altos incluso mayores que la regresión Múltiple, su uso está condicionado por la cantidad de ensayos de laboratorio para poder entrenarlas y probarlas.
- ✓ Es recomendable dominar las dos técnicas (análisis multivariado y redes neuronales) porque debido a las condiciones de tiempo, exactitud requerida, cantidad de observaciones es factible utilizar un método más que el otro. Por ejemplo, si se compara los datos del miembro Galembó, se obtuvo un  $R^2$  de 0.57 y 0.64, usando el análisis multivariado y las redes, respectivamente. La diferencia es de un 11%, a favor de la redes Neuronales, pero efectivamente involucran un costo mayor (ver tabla comparativa)

**Tabla 7-1.** Cuadro Comparativo de requerimientos entre las técnicas de análisis Multivariado y Redes Neuronales.

Requerimientos	Redes Neuronales Artificiales	Análisis Multivariado
Costo licencias (personal)	\$5000000 (matlab)	\$109.999 (office)
Personal	Conocimientos en programación-algoritmos genéticos	Conocimiento en office (Excel)
Tiempo para obtener resultados.	1 semana	15 minutos

✓ Los resultados mostraron que el registro Gamma Ray, es el segundo que más aporta a la correlación. El registro Gamma Ray, determina el volumen de arcillosidad de la formación, por tanto, está relacionado con la mineralogía. Las arcillas presentan alta plasticidad que las hace más deformables que otros minerales como el cuarzo; afectando directamente el módulo de Young, por tanto tener en cuenta el registro Gamma Ray en la predicción es concordante con la literatura. (ver capítulo 1).

✓ El módulo de Young está afectado por la porosidad total de la roca debido a que los poros también sufren deformación cuando se aplica una carga, en efecto los resultados de este trabajo mostraron que el registro de porosidad debe ser tenido en cuenta para predecir el módulo de Young.

✓ El registro resistivo que está relacionado con la saturación de agua en la roca, resultó no aportar en la correlación para el cálculo del módulo de Young en este caso específico del miembro Galembó de la formación la Luna en la orientación horizontal. Estos resultados pueden estar asociados por poca variación en la saturación de agua, que no es contradictorio con que el contenido de agua afecte el módulo de Young.

✓ El tiempo de transito de la onda de corte depende de un conjunto de parámetros que afectan el módulo de Young como son la porosidad, mineralogía, grado de alineación, presencia de planos de debilidad, fracturas. Los resultados de este estudio mostraron que el tiempo de transito de la onda de corte es la variable que aporta con más fuerza a la correlación para predecir el módulo de Young calculado del laboratorio a partir de registros eléctricos.

## 8 RECOMENDACIONES

- ✓ Continuar con el estudio de otras técnicas de análisis multivariado, como es el análisis discriminante, análisis clúster y componentes principales.
  
- ✓ Automatizar la herramienta para el cálculo con las redes Neuronales en cuanto a la determinación de los conjuntos de datos en entrenamiento y prueba.
  
- ✓ Estudiar la técnica conocida como máquina de soporte vectorial para conocer qué servicio pueden prestar y qué ventajas tiene sobre las redes y el análisis multivariado.
  
- ✓ Tener en cuenta el costo de la técnica de análisis multivariado con respecto a las redes Neuronales, en el caso que a nivel empresarial se necesite invertir en modelos predictivos.

## BIBLIOGRAFÍA

ALONSO, Luis y CALONGE, Teodoro. Capítulo 1 Redes Neuronales y Reconocimiento de patrones. En ALONSO, L. Reconocimiento de Patrones con Redes Neuronales .Salamanca: Imprenta cardenal, 2001. P. 1.1.

American Society for Testing and Materials: Standard Test Method For Compressive Strength And Elastic Moduli Of Intact Rock Core Specimens Under Varying States Of Stress And Temperatures. ASTM D7012, united States, 2010.

DEERE DU, Miller RP. Engineering Classification and Index Properties for Intact Rock. Department of Civil Engineering, University of Illinois, Urbana, IL 1966, pp. 90

DEHGHAN, S; SATTARI, Gh; CHEHREH, Chelganis y ALIABADI Ma. Prediction of uniaxial compressive strength and modulus of elasticity for travertine sample using regression and artificial neural networks. Mining Science and technology 20 (2010) 0041-0046

DEHGHAN, S. SATTARI, GH. CHEHREH Chelganis. “ Prediction of Uniaxial Compressive Strength and Modulus of Elasticity for Travertine Sample Using Regression and Artificial Neural Networks. Mining Science and Technology 0041-0046, 2009

GARAGON, Mert y CAN Tolga. Predicting the Strength anisotropy in uniaxial compression of some laminated Sandstone using multivariate regression analysis. Materials and structures (2010) 43: 509-5017.

HEMPKINS, W. Multivariate Statistical Approaches in Formation Evaluation. Standard oil Company of California. SPWLA EIGHTEENTH Annual Loguin Symposium, JUN 5-8, 1977. P 1-4

KONKAGUL, Engin y SANTI, Paul. Predicting the unconfined compressive strength of the Breathitt Shales using Slake durability, shores hardness and rock structural properties. International Journal of Rock Mechanics and Mining Sciences 36 (1999) 139-153

NEWBOLD, Paul. Estadística para Administración y Economía. 6 ed. Madrid: Pearson Educación, 2008. P. 494

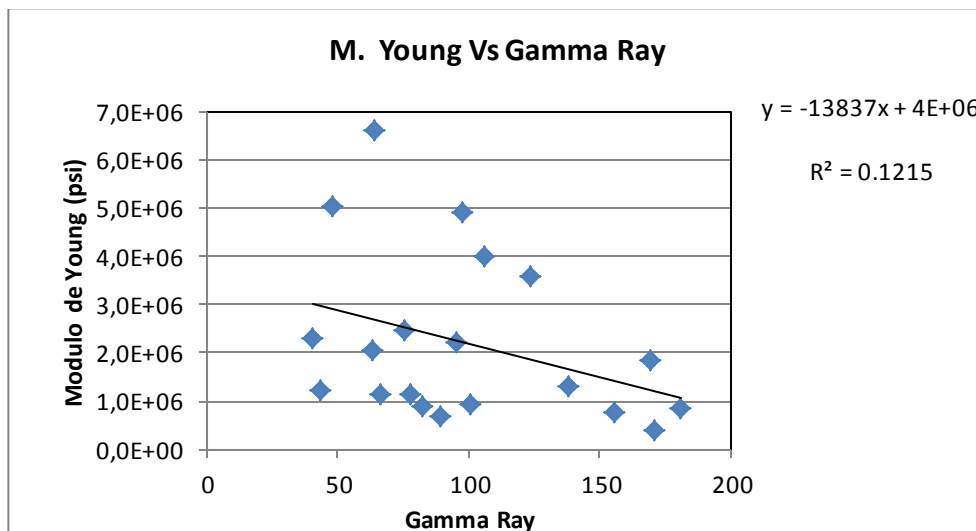
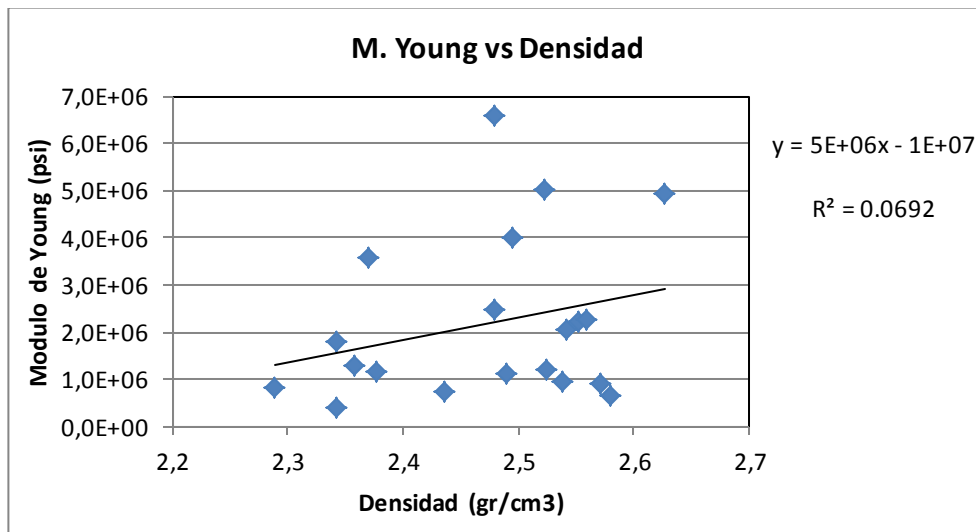
SINGH Rajesh, KAINTHOLA Ashutosh, SINGH T. N; Estimation of elastic constant of rocks using an ANFIS approach, department of Earth sciences , indian institute of technology Bombay; India. Elsevier: Septiembre de 2011.

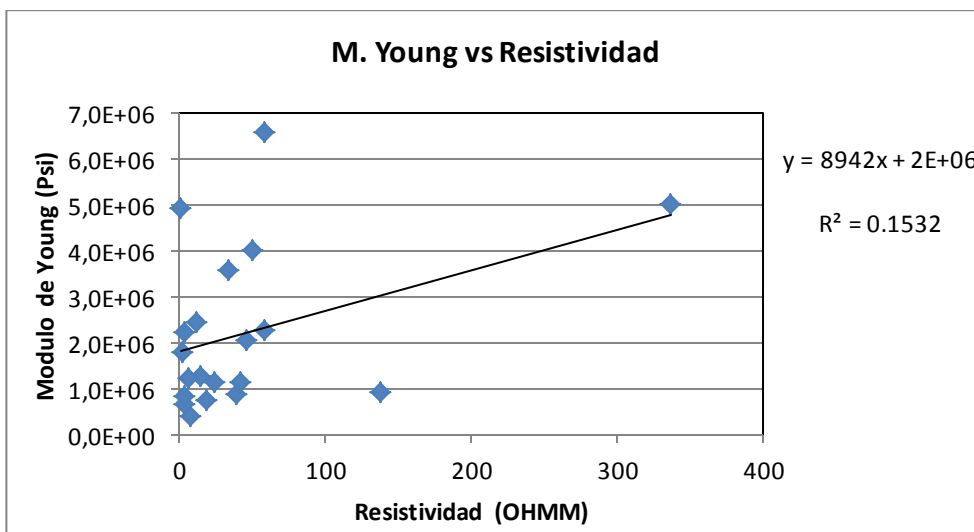
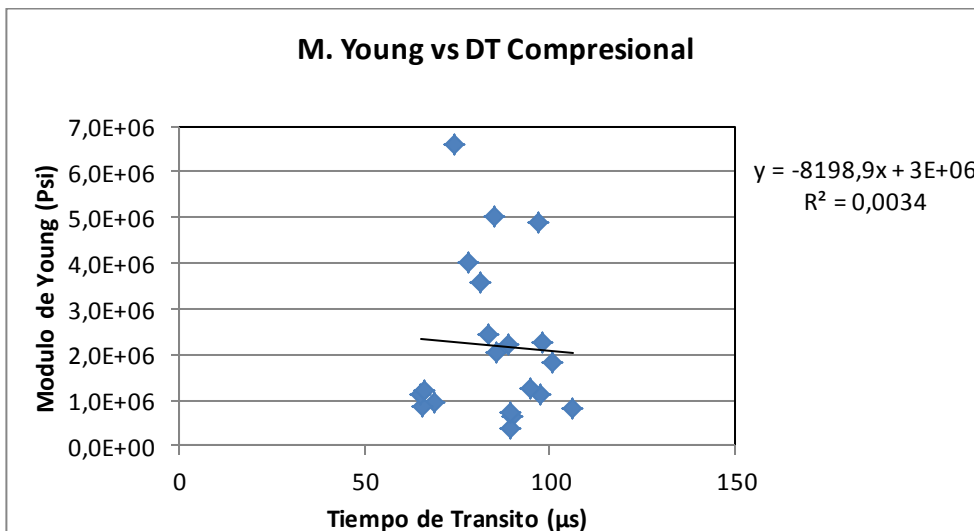
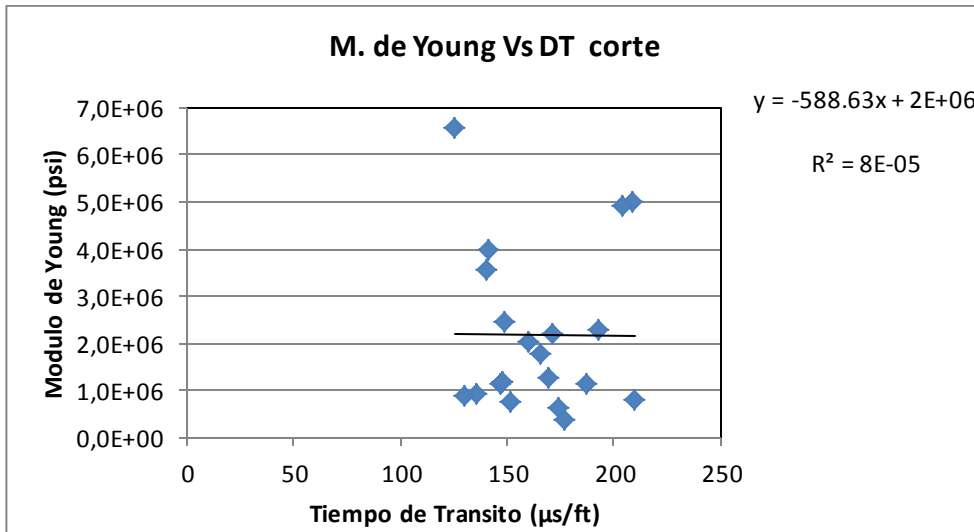
VIKAS Kumar, Tesis: Master en ciencias : “Geomechanical characterization of shale using Nano-indentation” , Universidad de Oklahoma, Norman: Oklahoma

# ANEXOS

## Anexo A

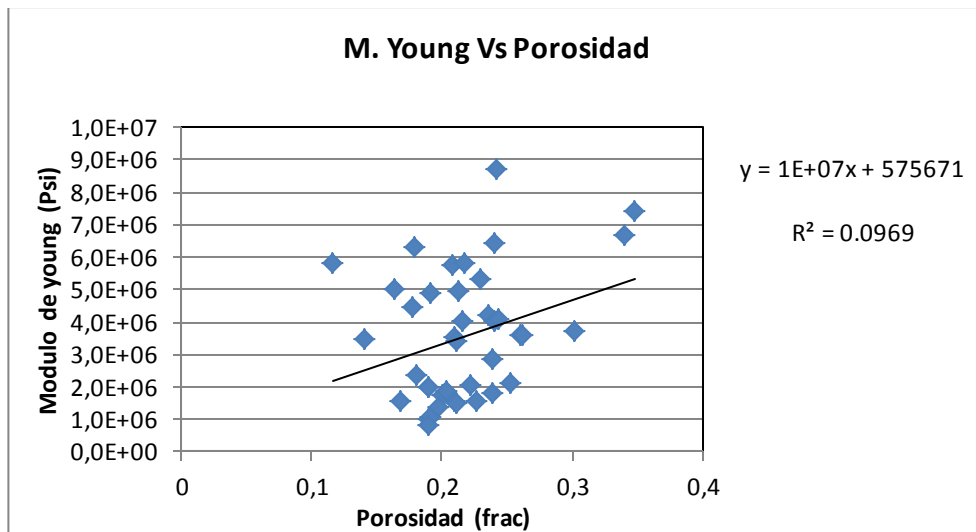
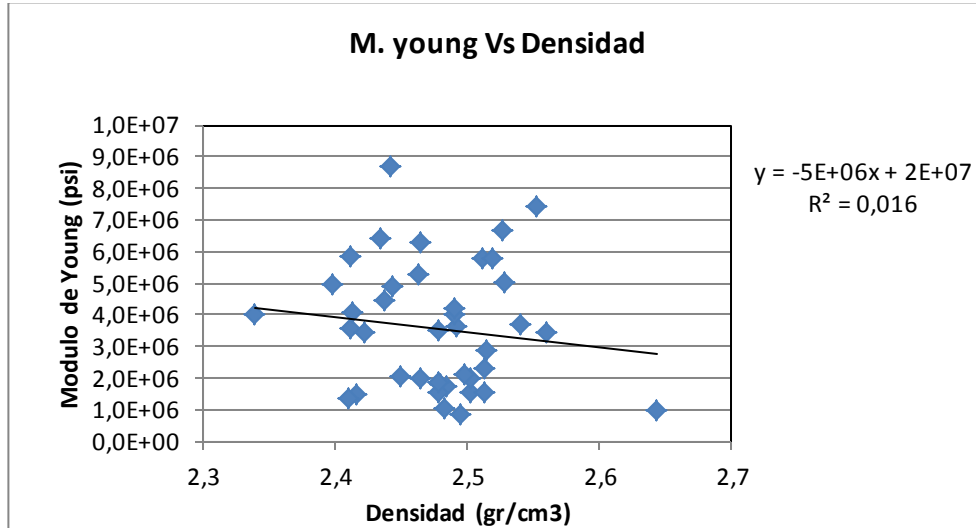
Graficas Regresión Simple Pozo La luna-1 Vertical. Formación la Luna, Simití y tablazo.

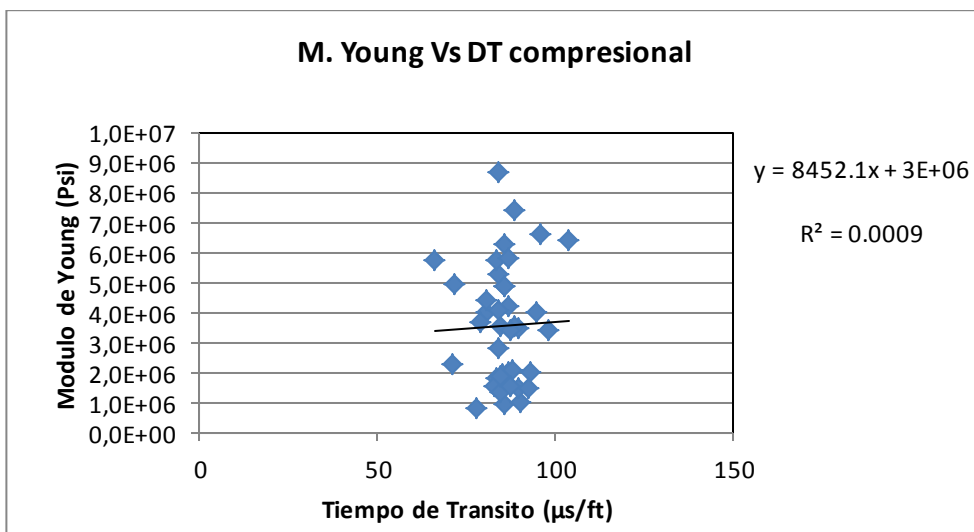
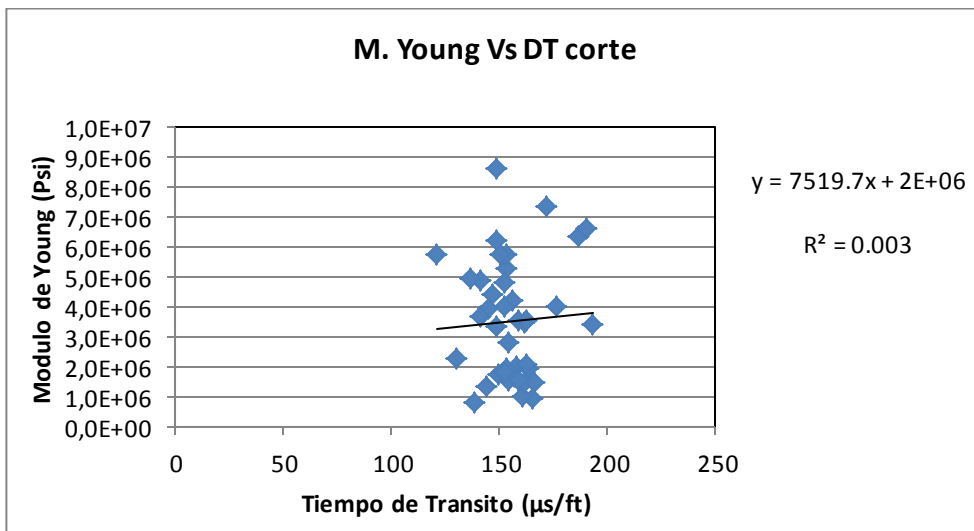
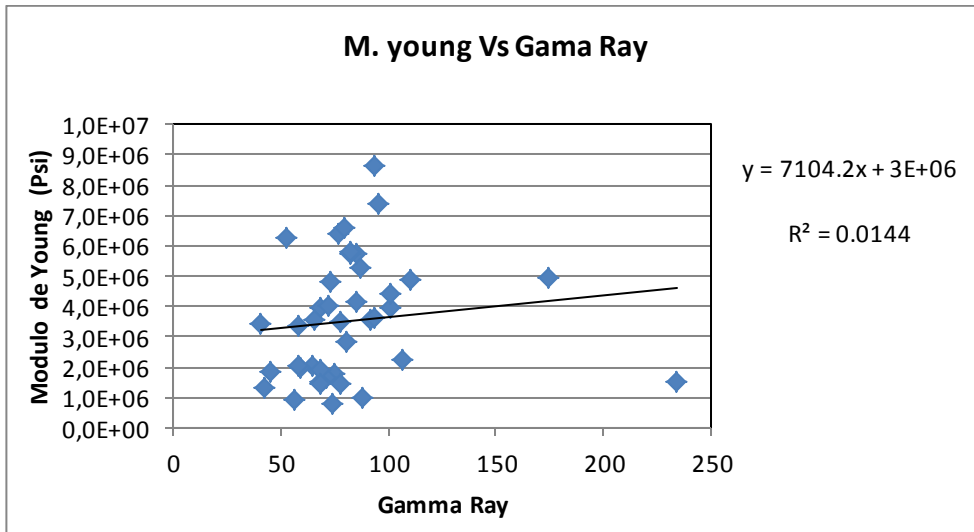


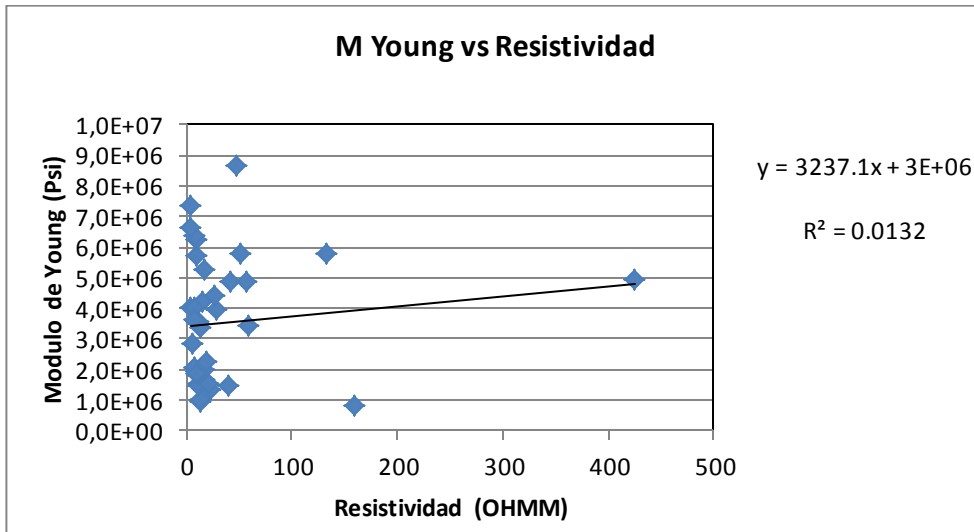


Anexo B

Graficas Regresión simple Formación La Luna del pozo la luna-1, Horizontal

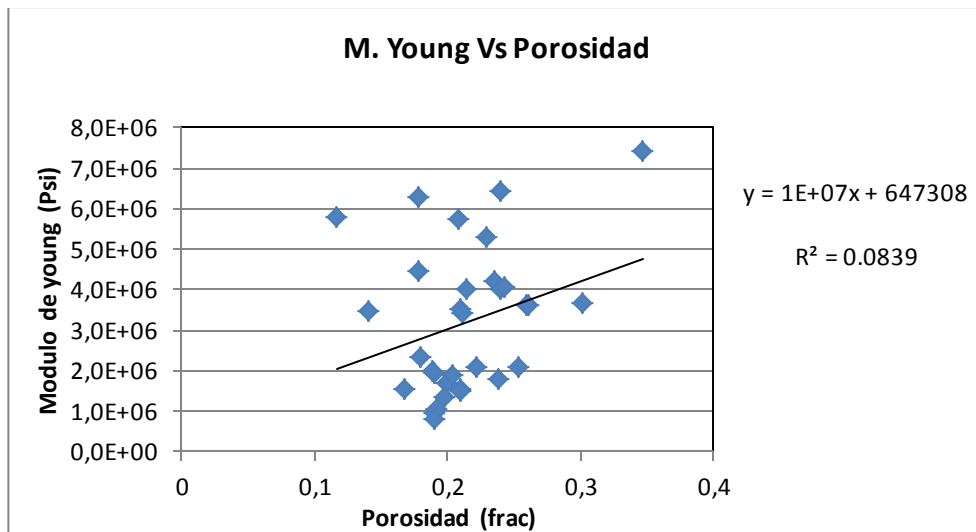
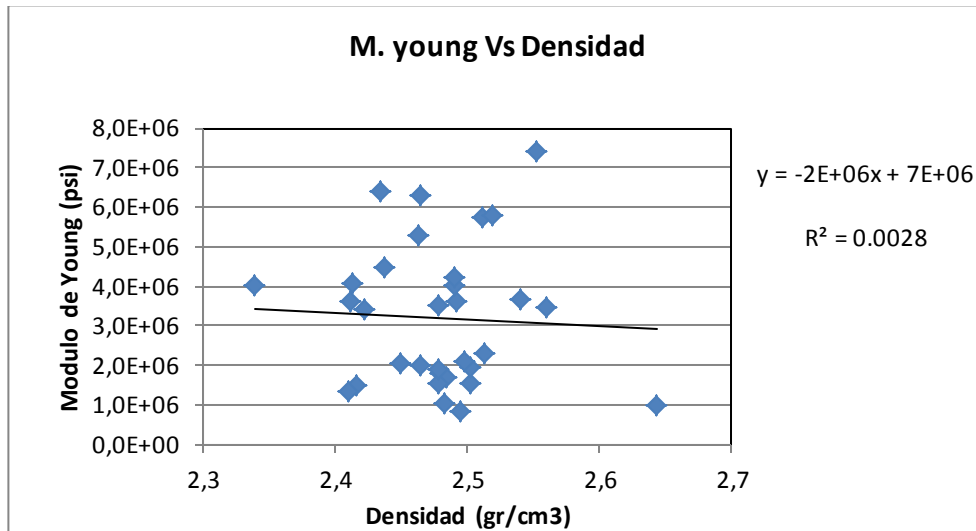


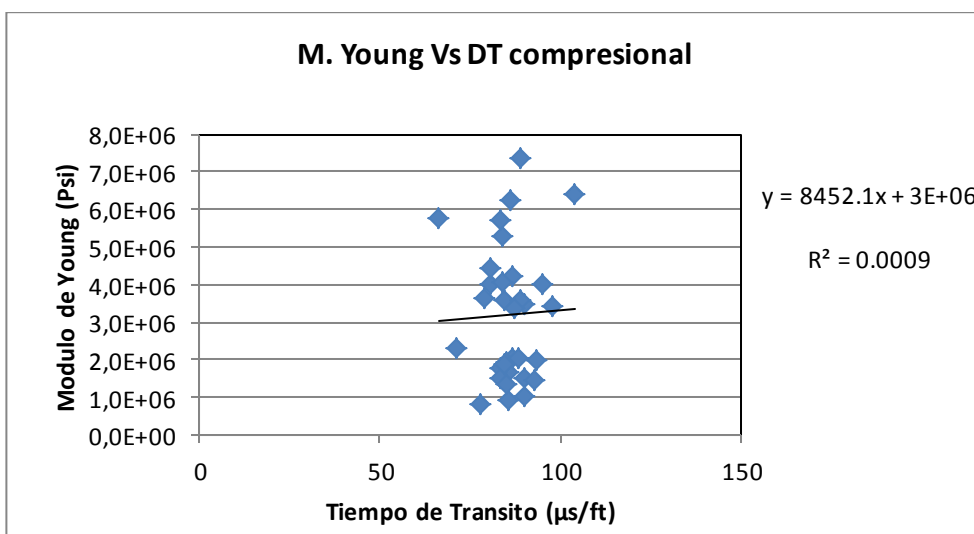
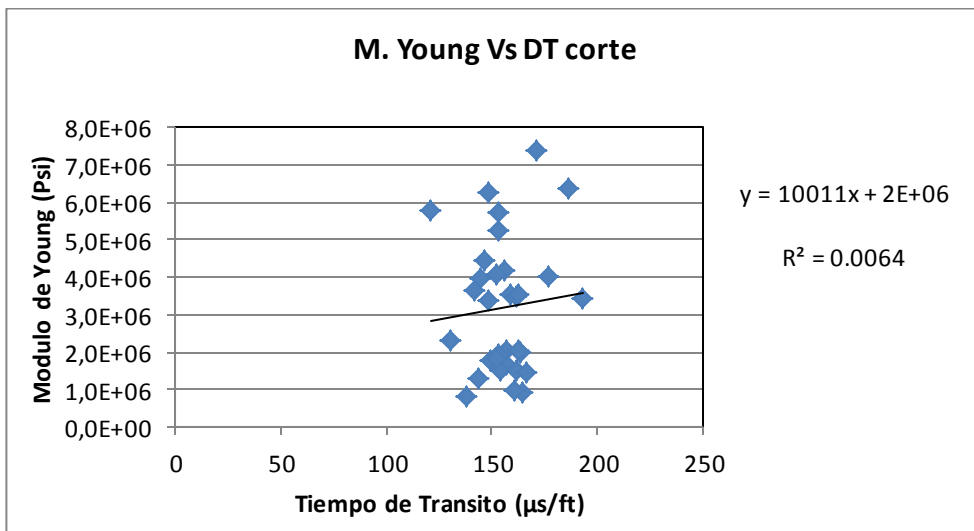
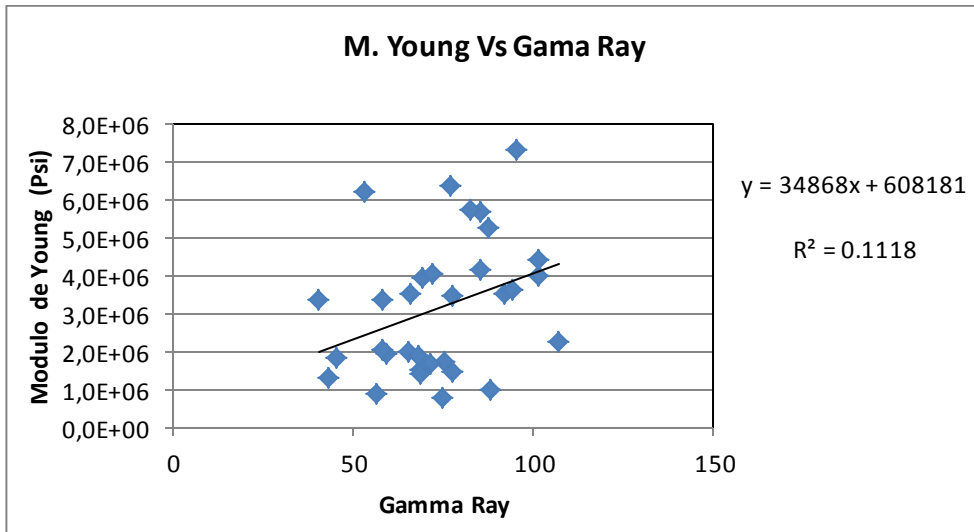


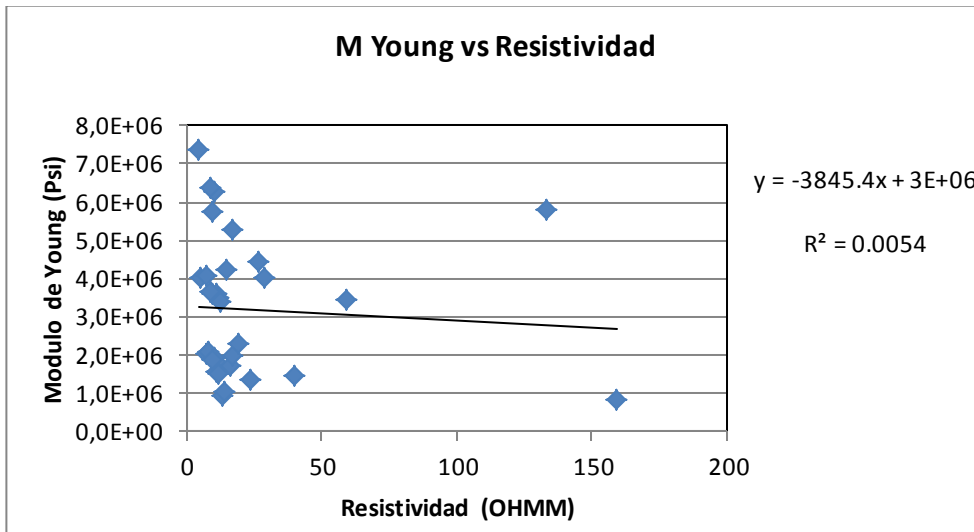


ANEXO C

Graficas Regresión Simple Formación la luna miembro Galembó Horizontal.







## ANEXO D

Graficas Regresión Simple Formación La Luna, miembro Galembro, Orientación Vertical

