



**FACULTAD DE INGENIERÍA ARQUITECTURA  
Y URBANISMO**

**ESCUELA ACADÉMICA PROFESIONAL DE INGENIERÍA  
DE SISTEMAS**

**TESIS**

**COMPARACIÓN DE LAS TÉCNICAS DE  
REGRESIÓN LINEAL MULTIPLE Y RED  
NEURONAL ARTIFICIAL PARA ESTIMAR EL  
GRADO DE SALINIDAD EN SUELOS CON  
ABUNDANTE VEGETACIÓN MEDIANTE EL  
PROCESAMIENTO DE IMÁGENES**

**PARA OPTAR TÍTULO PROFESIONAL DE INGENIERO  
DE SISTEMAS**

**Autor(a):**

**Bach. Sánchez Elescano Cesar Manuel**

**ORCID: <https://orcid.org/0000-0001-8550-6155>**

**Asesor(a):**

**Mg. Mejía Cabrera Heber Iván**

**ORCID: <https://orcid.org/0000-0002-0007-0928>**

**Línea de Investigación:**

**Infraestructura, Tecnología y Medio Ambiente**

**Pimentel – Perú 2020**

## Resumen

Las técnicas Regresión Lineal Múltiple y Red Neuronal Artificial son dos de las técnicas más utilizadas en los trabajos relacionados al presente trabajo, con mayor uso en la clasificación, estimación y predicción de datos a partir de varias variables explicativas y una variable esperada, no obstante, pocas investigaciones concluyen que técnica es la que entrega mejores resultados en cuanto a precisión. La presente tesis tiene como objetivo determinar que técnica es la más precisa para estimar salinidad en suelos con Abundante Vegetación (Cosecha de Arroz), las técnicas en estudio son: La regresión Lineal Múltiple y Red Neuronal Multicapa. Los datos de entrada para ambas técnicas son valores numéricos relacionados a un pixel de imágenes de tipo .TIFF obtenidas del procesamiento de imágenes del satélite Landsat 8, el cálculo de los valores de las imágenes nos da como resultados índices espectrales expresadas en una nueva imagen de tipo .TIFF, el pixel es ubicado dentro de las imágenes con respecto a una coordenada UTM SUR mapeada al obtener muestras de suelos, las cuales se procesaron en el balotario de la Universidad Pedro Ruiz Gallo para obtener valores numéricos de Conductividad Eléctrica. Utilizaron el NDVI, SAVI, VSSI, NDSI como índices espectrales y otros índices de Salinidad utilizados en estudios anteriores S1, S2, S3, S4, S5. Para la estimación de salinidad a partir de los valores numéricos de índices espectrales se utilizó un modelo de regresión expresado en ecuación resultante de la aplicación de la regresión lineal múltiple, también se utilizó la función de red neuronal de la herramienta Weka, ambos resultados fueron comparados. La variable utilizada para la determinación es el Error Medio Absoluto y Error Cuadrático Medio. Los resultados muestran un mejor desempeño de la técnica de Red Neuronal Artificial al estimar salinidad en suelos con abundante vegetación con un Error Cuadrático de 0.1753.

Se recomienda para futuros estudios, utilizar una cantidad mucho mayor de muestras de suelo (CE), para ajustar los valores resultantes de estimación en cada técnica.

**Palabras Clave:** Regresión Lineal Múltiple, Red Neuronal Artificial, Salinidad, abundante vegetación, Imágenes Multiespectrales, Índices espectrales, Satélite Landsat 8, Error Cuadrático medio

## **Abstract**

The Multiple Linear Regression and Artificial Neural Network techniques are two of the most used techniques in the works related to this work, with greater use in the classification, estimation and prediction of data from several explanatory variables and an expected variable, however, few investigations conclude that the technique is the one that provides the best results in terms of precision. The objective of this thesis is to determine which technique is the most accurate to estimate salinity in soils with Abundant Vegetation (Rice Harvest), the techniques under study are: Multiple Linear Regression and Multilayer Neural Network. The input data for both techniques are numerical values related to a pixel of .TIFF type images obtained from the image processing of the Landsat 8 satellite, the calculation of the values of the images gives us spectral indices as results expressed in a new Image of type .TIFF, the pixel is located within the images with respect to a UTM SUR coordinate mapped when obtaining soil samples, which were processed in the ballotary of the Pedro Ruiz Gallo University to obtain numerical values of Electrical Conductivity. They used the NDVI, SAVI, VSSI, NDSI as spectral indices and other salinity indices used in previous studies S1, S2, S3, S4, S5. For the estimation of salinity from the numerical values of spectral indices, a regression model expressed in an equation resulting from the application of multiple linear regression was used, the neural network function of the Weka tool was also used, both results were compared. The variable used for the determination is the Absolute Mean Error and Mean Square Error. The results show a better performance of the Artificial Neural Network technique when estimating salinity in soils with abundant vegetation with a Square Error of 0.1753.

It is recommended for future studies to use a larger quantity of soil samples (CE), to adjust the resulting estimation values in each technique.

**Keywords:** Multiple Linear Regression, Artificial Neural Network, Salinity, Abundant Vegetation, Multispectral Images, Spectral Indexes, Landsat Satellite 8, Mean Quadratic Error.