

UNIVERSIDAD NACIONAL DE INGENIERÍA
FACULTAD DE CIENCIAS



TESIS

**“REDES NEURONALES ARTIFICIALES APLICADAS A LA
REMOCIÓN DE COBRE EN AGUA UTILIZANDO
QUITOSANO ENTRECruzADO CON
GLUTARALDEHIDO”**

PARA OPTAR EL GRADO ACADEMICO DE MAESTRO EN CIENCIAS
CON MENCIÓN EN QUÍMICA

ELABORADA POR:

MANUEL ENRIQUE BLANCO LECCA

ASESOR:

Mg. CHRISTIAN RONALD JACINTO HERNÁNDEZ

LIMA – PERÚ

2018

A mi amada familia: la cual no ha cesado incansablemente, desde el día que nací, en darme el empuje de perseverancia y deseo de superación. Todo cuando he logrado es gracias a ustedes, y hoy especialmente, están presentes en este trabajo de tesis.

Este trabajo es ciertamente un regalo de amor para ustedes, familia linda, y pese a que hoy me encuentro lejos de ustedes, nuestro amor nos une venciendo diferencias de tiempo y espacio.

¡Los amo!

AGRADECIMIENTOS

Deseo agradecer de una manera muy especial a todas las personas que me dieron aliento para poder realizar este trabajo de investigación: a mis familiares, amigos, compañeros de universidad, profesores, etc. Quiero que sepan que toda esa inyección de fuerza que me han entregado ha sido un factor muy significativo para terminar la presente tesis.

Agradezco de manera especial al personal de servicio del Laboratorio por el apoyo amable en los materiales y reactivos que se requirieron para el desarrollo experimental de este trabajo, a los profesores: Ily Maza, Rosario López, Samuel Oporto, por todas las observaciones, la visión crítica y por los aportes a este trabajo y a las ayudas brindadas en las consultas realizadas, las cuales fueron absorbidas con mucha disponibilidad e interés por los docentes citados.

Agradezco a su vez a ALS Life Sciences Perú, en su sede de Lima, por las ayudas en las determinaciones cuantitativas, cediéndome el espacio para el uso del Espectrofotómetro de Absorción Atómica y todos los materiales y muestras de aguas que me brindaron.

Finalmente, quiero agradecer a al profesor Christian Jacinto, asesor de la presente tesis, por creer en este proyecto, y todo el gran apoyo brindado tanto en la parte teórica como en la experimental, en la dedicación y la paciencia hacia mi persona y en el interés y la confianza depositada en este trabajo de investigación.

ÍNDICE

AGRADECIMIENTOS.....	ii
ÍNDICE.....	iii
RESUMEN TABLAS	vi
Tabla resumen de “Tablas”:	vi
Tabla resumen de “Gráficos”:	vii
Tabla de Abreviaturas	viii
RESUMEN	1
ABSTRACT	2
INTRODUCCIÓN.....	3
ESTADO DEL ARTE	7
OBJETIVOS.....	10
OBJETIVOS GENERALES.....	10
OBJETIVOS ESPECÍFICOS	10
MARCO TEÓRICO	11
1. QUITOSANO	11
1.1 Breve reseña histórica del Quitosano:.....	12
1.2 Química del quitosano:.....	14
1.2.1 Reacciones de entrecruzamiento:	14
2. EL COBRE	16
2.1 Química del Cobre	16
2.2 El cobre como contaminante	16
2.3 Reseña toxicológica.....	18
2.4 Valor máximo permisible de cobre en el Perú (Ministerio del Ambiente, 2017) ..	19
3. BIOSORCIÓN	20
3.1 Biosorción con quitosano	21
3.2 Propiedades de acomplejación	22
4. REDES NEURONALES ARTIFICIALES	24
4.1 Introducción	24
4.2 Definición.....	24
4.2.1 Descripción de las RNA	24
4.2.2 Comparación con otra técnica de Inteligencia Artificial	24

4.3 Idea de red neuronal artificial.....	27
4.3.1 Redes Neuronales Artificiales (RNA).....	27
4.3.2 Redes Neuronales Supervisadas y No Supervisadas	30
a. Reglas de entrenamiento Supervisada	30
b. Reglas de Entrenamiento No Supervisada.....	31
4.3.3 Funciones de Base y de Activación.....	32
a. Tipos de Función Base	32
b. Función de Activación.....	33
4.3.4 Estructuras de conexión de atrás hacia delante	34
4.3.5 Tamaño de las Redes Neuronales.....	36
4.4 Métodos de regresión y redes neuronales	36
4.5 Perceptrón Simple	38
4.5.1 Limitaciones del perceptrón	40
4.6 Perceptrón Multicapas.....	41
4.6.1 Arquitectura	42
4.6.2 Algoritmo backpropagation.....	43
4.6.3 Etapa de funcionamiento	43
4.6.4 Etapa de aprendizaje.....	44
4.6.5 Fases en la aplicación de un perceptrón multicapa.....	46
4.6.6 Selección de las variables relevantes y preprocesamiento de los datos	46
4.6.7 Entrenamiento de una neurona	47
a. Elección de los pesos iniciales	47
b. Arquitectura de la red	47
4.6.8 Tasa de aprendizaje y factor momento	48
4.6.9 Función de activación de las neuronas ocultas y de salida.....	49
4.6.10 Evaluación del rendimiento del modelo	49
4.6.11 Interpretación de los pesos obtenidos.....	50
PARTE EXPERIMENTAL.....	51
1. EQUIPOS Y REACTIVOS	52
1.1 Reactivos	52
1.2 Equipos.....	52
2. PROCEDIMIENTO EXPERIMENTAL	52
2.1 Preparación de quitosano 5%	52

2.2 Preparación de hidróxido de sodio 2.5N	53
2.3 Preparación de perlas de quitosano	53
2.4 Entrecruzamiento de quitosano con glutaraldehído	53
2.5 Preparación de solución de cobre 100 ppm.....	54
2.6 Procedimiento experimental.....	54
2.7 Definición de la red neuronal artificial	57
RESULTADOS OBTENIDOS Y DISCUSIONES.....	58
CONCLUSIONES.....	88
CONCLUSIONES	88
NUEVOS TRABAJOS DE INTERÉS PARA EL LECTOR	90
REFERENCIAS BIBLIOGRÁFICAS	91
ANEXO	96
Algoritmo: redes neuronales artificiales multicapa con aprendizaje backpropagation ...	97
Constancia de participación de la tesis de maestría en Congreso Peruano de Química	100

RESUMEN TABLAS

Tabla resumen de “Tablas”:

Tabla 1: Propiedades químicas del Cu en el ambiente.....	16
Tabla 2: Datos de referencia del Cu como contaminante. (OMS, Cooper in drinking-water, 2003)	17
Tabla 3: Límite Máximo Permisible de Cobre según legislación peruana (MINAM) - ECA aguas 2017.....	19
Tabla 4: Comparación entre redes neuronales artificiales y sistemas expertos (Slater, Hazen, & Sakthivel, 1993).	25
Tabla 5: Algoritmos de aprendizaje para tareas de análisis de datos.	29
Tabla 6: Funciones de activación más comúnmente usadas.	33
Tabla 7: Tabla de niveles experimentales	54
Tabla 8: Matriz de diseño de experimentos	55
Tabla 9: Valores fijados en la elaboración de la RNA.....	59
Tabla 10: MSE determinado por cada RNA, con diferentes números de capas ocultas (por WEKA v.3.6).	59
Tabla 11: Resultados obtenidos de la RNA entrenada, con 16 capas ocultas y con los parámetros fijados en la Tabla 9 y Gráfico 18	65
Tabla 12: Errores máximos determinados comparando el %Rpredictivo vs %Rexperimental, para la RNA de la Tabla 11	67
Tabla 13: Resultados obtenidos de la RNA previamente entrenada, con 16 capas ocultas y con los parámetros fijados para la validación.	73
Tabla 14: Resultados obtenidos de la RNA previamente entrenada con la matriz expandida, con 16 capas ocultas.....	73
Tabla 15: Comparación de los resultados obtenidos con la RNA entrenada con la matriz del DOE y la entrenada con la matriz expandida (incluyendo patrones de validación).	80
Tabla 16: Resultados de optimización (%R al 100% de Cu) en función de la masa de quitosano / GLA, en nivel ácido, neutro y básico.	82
Tabla 17: Comparación de las predicciones del RNA, DOE 2 ⁶ (+5PC) y regresión lineal múltiple.	82
Tabla 18: Errores máximos y error cuadrático medio de tres modelos de predicción: RNA, DOE, regresión lineal múltiple.....	84

Tabla resumen de “Gráficos”:

Gráfico 1: Unidad repetitiva del quitosano.	13
Gráfico 2: Mecanismo de reacción de la desacetilación de la quitina para obtener quitosano. a) Reacción generalizada, b) Detalle del mecanismo. (Marinela, 2014)	13
Gráfico 3: Unidad repetida de quitano	13
Gráfico 4: Reacciones de entrecruzamiento para producir hidrogeles de quitosano usando un dialdehído como agente entrecruzante. (Díaz Sánchez, 2011).....	15
Gráfico 5: Modelos de acomplejación de metales tipo a) colgante y b) puente. (Lasheras Zubiate, 2011)	23
Gráfico 6: Similitud operativa entre una neurona biológica y una neurona artificial.	28
Gráfico 7: Esquema funcional del núcleo de la neurona artificial.	28
Gráfico 8: Esquema de sistema de entrenamiento supervisado.	31
Gráfico 9: Esquema de sistema de entrenamiento no supervisado.	31
Gráfico 10: Tipos de redes de pesos de conexión entre dos capas de neuronas.....	34
Gráfico 11: Red neuronal parcialmente interconectada.	35
Gráfico 12: Red neuronal de un input (ejemplo).....	37
Gráfico 13: Red neuronal de varios inputs (ejemplo).	37
Gráfico 14: Esquema del Perceptrón simple.	38
Gráfico 15: Margen de caracterización del perceptrón.	40
Gráfico 16: Esquema de un perceptrón multicapas.....	42
Gráfico 17: RNA con 16 capas ocultas (6 – 16 – 1). Gráfico proporcionado en WEKA.	60
Gráfico 18: RNA con 16 capas ocultas (6 – 16 – 1). Gráfico proporcionado en MATLAB.....	61
Gráfico 19: Parámetros para entrenamiento de la RNA (6 – 16 – 1).	61
Gráfico 20: Fases y resultados del entrenamiento de la RNA (6 – 16 – 1).	62
Gráfico 21: Perfil del rendimiento de la validación de la RNA (6 – 16 – 1).	62
Gráfico 22: Perfil de evolución de los parámetros de la RNA por épocas de aprendizaje.....	63
Gráfico 23: Gráfica de correlación en la fase de entrenamiento, validación, prueba.....	64
Gráfico 24: Gráfica de dispersión de los valores de %R predictivo vs experimental	68
Gráfico 25: Gráfico QA/QC de la regresión %R predictiva vs %R experimental dada por la matriz sintética (agua desionizada). Gráfica hecha por Minitab ® 17.	71
Gráfico 26: Análisis residual de la regresión %R predictiva vs %R experimental dada por la matriz sintética (agua desionizada). Gráfica hecha por Minitab ® 17.	72
Gráfico 27: Gráfica de dispersión de los valores de %R predictivo vs experimental en la matriz de experimento expandida.	76
Gráfico 28: Gráfico QA/QC de la regresión %R predictiva vs %R experimental dada por la matriz expandida. Gráfica hecha por Minitab ® 17.	79
Gráfico 29: Análisis residual de la regresión %R predictiva vs %R experimental dada por la matriz expandida. Gráfica hecha por Minitab ® 17.	80
Gráfico 30: Perlas de quitosano entrecruzadas con glutaraldehído.....	81
Gráfico 31: Gráfico de Pareto del análisis del DOE 2 ⁶ (+5PC)	85
Gráfico 32: Gráfica de los efectos de los factores principales para el %Remoción	86

Tabla de Abreviaturas

<i>RNA</i>	Redes Neuronales Artificiales
<i>ANN</i>	Artificial Neural Network
<i>MSE</i>	Mean Squared Error
<i>SMEWW</i>	Standard Methods for the Examination of Water and Wastewater
<i>GLA</i>	Glutaraldehído
<i>QA/QC</i>	Quality Assurance & Quality Control
<i>DoE</i>	Design of Experiments
<i>ANOVA</i>	Analysis of Variance
<i>PC</i>	Puntos centrales
<i>LMP</i>	Límite Máximo Permisible
<i>RLM</i>	Regresión Lineal Múltiple
<i>OMS</i>	Organización Mundial de la Salud

RESUMEN

Con el objeto de estudiar un sistema inteligente capaz de predecir el comportamiento porcentual en la remoción de cobre en aguas desionizada, se ha diseñado una técnica de inteligencia artificial aplicado al problema químico de remoción de Cu (II) en aguas, teniendo como soporte de programación el software MATLAB R2014a. Este sistema ha sido entrenado inicialmente con 74 datos ingresados como patrones de aprendizaje, teniendo como parámetros de ingreso para la red neuronal artificial 6 parámetros críticos de la experimentación: pH, masa de quitosano entrecruzado con glutaraldehído (g), concentración de glutaraldehído (GLA) (g/100mL), tiempo de agitación (min), frecuencia de agitación (rpm), y concentración inicial de cobre (ppm). Se verificó que el máximo error de predicción del modelo de red neuronal artificial del tipo perceptrón multicapa (14.15%) se obtiene a partir de 16 capas ocultas. Para comprobar la efectividad de la red neuronal, se hizo uso de 9 mediciones distintas a las de la matriz de experimentos, para así realizar la validación de este sistema, verificando así la efectividad de nuestra red neuronal (a un 98.79% de correspondencia entre los valores predictivos, respecto a los valores reales observados). La descontaminación se logra por captura del ion Cu (II) (analito de estudio en aguas) por las perlas de quitosano entrecruzado con glutaraldehído. La obtención de los datos experimentales se toma por diferencia porcentual entre la concentración inicial del agua (contaminada con $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$) y la concentración obtenida después del proceso de descontaminación con las perlas de quitosano entrecruzadas con glutaraldehído, registrado por la técnica de espectrofotometría de absorción atómica.

ABSTRACT

To study an intelligent system capable of predicting the percentage behavior in the removal of copper in deionized waters, an artificial intelligence technique has been designed applied to the chemical problem of metal removal (specifically Cu (II)) in waters, taking MATLAB R2014a as software for to programming support. This system has been initially trained with 74 data entered as learning patterns, having as input parameters, for the artificial neural network, 6 critical parameters of the experiment: pH, mass of chitosan cross-linked with glutaraldehyde (g), concentration of glutaraldehyde (GLA) (g / 100mL), stirring time (min), stirring frequency (rpm), and initial copper concentration (ppm). It was verified that the maximum prediction error of the artificial neural network model of the multilayer perceptron type (14.15%) is obtained from 16 hidden layers. To check the effectiveness of the neural network, 9 different measurements were used to those of the experiment matrix, in order to validate this system, thus verifying the effectiveness of our neural network (at 98.79% correspondence between the predictive values, with respect to the actual values observed). Decontamination is achieved by capturing the Cu (II) ion (analyte studied in water) by the chitosan beads cross-linked with glutaraldehyde. Obtaining the experimental data is taken as a percentage difference between the initial water concentration (contaminated with $\text{CuSO}_4 \cdot 5\text{H}_2\text{O}$) and the concentration obtained after the decontamination process with the chitosan beads cross-linked with glutaraldehyde, registered by the absorption spectrophotometry technique atomic.

INTRODUCCIÓN

El cobre es un metal de transición ampliamente usado en la industria para la manufactura de muchos productos tales como amalgamas con diversas aplicaciones, agroquímicos (especialmente fungicidas y micostáticos), esmaltes y pigmentos, reactivos para curtiembre, alguicidas de uso profuso en cuerpos de agua dulce, prótesis médico quirúrgicas y dispositivos intrauterinos (DIU), entre otros. Por este motivo, no es sorprendente que el hombre contamine el medio ambiente con progresivas cantidades de cobre. La contaminación con este metal se produce mayoritariamente en la cercanía de las minas donde se extrae y procesa, y en las refinerías durante su purificación (Riedel, 2008).

Los polvos de Cu que se liberan al aire y al agua reaccionan con el oxígeno y el azufre, que muchas veces son parte de los mismos minerales desde donde se refina el metal, generando compuestos aún más hidrosolubles y diseminables. El agua superficial que se contamina de esta manera puede drenar hacia los acuíferos y aumentar el nivel de cobre en ellos (Riedel, 2008) y por consiguiente en las redes de agua potable domiciliarias (Sharunda D. Buchanan, 1994).

Las zonas urbanas también son sitios de gran contaminación con Cobre. Esto se debe, además de la cercanía con las fábricas, al uso profuso de combustibles fósiles. En las zonas rurales en cambio, la principal fuente de contaminación con cobre la constituye el uso excesivo o inapropiado de pesticidas, fungicidas y alguicidas. Si bien se sabe que mueren alrededor de 300 000 personas por año en zonas rurales por causa de la exposición aguda a pesticidas formulados a base de cobre, es frecuente que se produzcan intoxicaciones por exposiciones crónicas debidas al consumo diario de vegetales que han sido tratados con agroquímicos (Gunay N, 2006). En áreas industrializadas el problema es mayoritariamente la exposición crónica a dosis bajas (Eddleston M, 2007). Los casos de exposición aguda se

producen sobre todo por la ingesta de líquidos contaminados con cobre, o por el consumo accidental o ex profeso de grandes cantidades de sales de cobre (Uriu-Adams JY, 2005).

También se pueden encontrar grandes concentraciones de cobre en peces u otros animales marinos. Por ejemplo, en las ostras del río Thames (Inglaterra) y de la bahía de Chesapeake (Taiwan) se puede observar un color verde en la carne del molusco debido a la deposición de sales cúpricas, condición que causa intoxicación aguda en las personas que los consumen (Riedel, 2008). A su vez, Green-Ruiz y Páez-Osuna (2001) encontraron que en la ensenada Altata ubicada en la laguna Culiacán (México) -en la cual se eliminan los desechos de la actividad agrícola local- existe una relación directa entre la concentración de Cr, Cu, Mn y Pb, y las descargas agrícolas registradas. En España, el cobre es el único compuesto aceptado como fungicida en la producción orgánica, y la Comisión de Regulación Europea permite un máximo de uso de 6 Kg/ha/año (*Commission Regulation - EC/473/2002*). Sin embargo, el umbral utilizado supera ampliamente lo establecido por esta comisión (Vicent A, 2009). De modo que el cobre puede ingresar a nuestro sistema por sobrecarga de la cadena nutricional a partir de suplementos para animales de cría, fórmulas comerciales para alimentación enteral o parenteral, o residuos de pesticidas y otros agroquímicos (Rotilio G, 2000).

La sorción es un método muy efectivo extensamente usado en la remoción de metales tóxicos desde contaminantes acuosos. El uso de biopolímeros como sorbentes es una técnica que apareció para el tratamiento de industrias que contienen efluentes de metales tóxicos, esto ha sido una investigación muy interesante (Quer-Brossa, 1983).

El biopolímero de quitosano resultó ser un efectivo sorbente de metales de transición. Esta capacidad es, sin embargo, influenciado por el pH de la solución. Varias modificaciones químicas han sido aplicadas al quitosano para mejorar la interacción con los iones metálicos

(Quer-Brossa, 1983) (Crescenzi, Imbriaco, Lárez, Dentini, & Ciferri, 1995) (Burriel, Conde, Jimeno, & Méndez, 1994)

El quitosano posee un valor del pKa ~ 5,6; siendo soluble solo en soluciones ácidas con valores de pH < 6, se ha venido buscando extender el intervalo de pH donde se pueda utilizar dicho biopolímero en solución, para darle una mayor aplicabilidad. En ese sentido, un material con una mayor densidad de carga generaría seguramente una mayor solubilidad; por ello, una de las rutas propuestas para lograr este objetivo es lograr una alta eficiencia en la cuaternización de los grupos aminos en las cadenas del polímero. Existen diversos métodos de cuaternización de grupos amino primario y secundario, algunos de los cuales ya se han aplicado con éxito al quitosano, obteniéndose diferentes derivados alquilados.

En las pasadas décadas, los químicos se han acostumbrado a utilizar ordenadores y, consecuentemente, a poner a punto diferentes métodos estadísticos complejos, tratando de explorar cada vez con más detalle las correlaciones multidimensionales entre variables predictoras y respuestas. Junto con la mejora de parámetros de calidad de los métodos analíticos tales como exactitud o precisión, quedó claro que todas las propiedades de interés no pueden describirse mediante relaciones unidimensionales o incluso relaciones multidimensionales lineales. Por tanto, los químicos analíticos desean siempre explorar las posibilidades de métodos nuevos que puedan ayudar a resolver estos problemas.

Uno de los métodos, o, para ser más precisos, un conjunto de métodos, que han despertado un interés general aplicándose con profusión en los últimos años son las redes neuronales artificiales. Desde la década de los 40, en la que nació y comenzó a desarrollarse la informática, el modelo neuronal la ha acompañado. De hecho, la aparición de los computadores digitales y el desarrollo de las teorías modernas acerca del aprendizaje y del procesamiento neuronal se produjeron aproximadamente al mismo tiempo, a finales de los años cuarenta.

Desde entonces hasta nuestros días, la investigación neurofisiológica y el estudio de sistemas neuronales artificiales (ANS, Artificial Neural Systems) han ido de la mano. Sin embargo, los modelos de ANS no se centran en la investigación neurológica, si no que toma conceptos e ideas del campo de las ciencias naturales para aplicarlos a la resolución de problemas pertenecientes a otras ramas de las ciencias y la ingeniería. Es difícil describir las RNAs con una definición simple. Quizás la descripción más acertada sea una comparación con una caja negra que posee múltiples entradas y múltiples salidas operando mediante un gran número de unidades aritméticas simples que se conectan entre sí en paralelo. El aspecto más importante a recordar de los métodos basados en RNAs es el hecho que funcionan mejor cuando tratan relaciones no lineales entre variables de entrada y salida. Las RNAs pueden emplearse también para describir o encontrar relaciones lineales pero el resultado final es a menudo peor que el encontrado mediante técnicas estadísticas más simples. Debido al hecho que al principio de la experimentación frecuentemente no se conoce si la relación entre respuesta y variables predictoras es o no lineal, un buen consejo consiste en aplicar siempre alguna técnica quimiométrica estándar para interpretar los datos junto con las RNAs (Blanco & Cerdà, 2007).

Es necesario resaltar la significación e importancia que las redes neuronales están adquiriendo en la actualidad como lo evidencia el hecho de formar parte de los estudios centrales de instituciones gubernamentales a nivel mundial. Así que la intención principal es profundizar en esta nueva tecnología, aprovechando el hecho de que será esta una materia en cursos posteriores en el transcurso de mi carrera, para así lograr un alto grado de aprendizaje y con el la implementación practica de alguna red neuronal. (Hilera & Martínez, 2000).

ESTADO DEL ARTE

Haciendo una revisión de los trabajos antecesores realizados al presente trabajo, se puede ver que es en estos últimos años, y particularmente en este 2018, que la técnica supervisada de redes neuronales artificiales ha mostrado ser de mucho interés de investigación.

Muchos trabajos dedicados a la predicción del porcentaje de remoción (%R) de metales en aguas prueban que las Redes Neuronales Artificiales (RNA) es ciertamente una herramienta confiable y de basta aplicación.

Sushovan Sen en la India aplicó las RNA como una herramienta predictiva, y el análisis de superficie de respuesta como herramienta de optimización, para la remoción de cromo (VI) en aguas residuales utilizando una biomasa de cianobacterias. El trabajo presentado consideró como variables de entrada la concentración inicial del Cromo (VI) en el agua a tratar, la temperatura a la cual se somete el proceso de sorción, la variación de pH y la cantidad de biomasa de cianobacteria. El trabajo citado confirmó la aplicación de la RNA en el proceso descrito. (Sushovan, Somnath, & Susmita, 2018)

Se aplicó también la RNA y un sistema de inferencia neuro-difusa (ANFIS) a la predicción de remoción de colorante (aserrín) y ion cobre (II) en soluciones acuosas, teniendo como variables de entrada la concentración inicial del colorante y del Cu(II), pH, tiempo de contacto con el sorbente y cantidad de sorbente utilizado. En este estudio también se analizó las propiedades cinéticas de la desorción. (Dolatabadi, Mehrabpour, Esfandyari, Alidadi, & Davoudi, 2018).

Las RNA, aplicadas al diseño de experimentos de Box-Behnken para el análisis predictivo de Cu(II) en aguas, teniendo como sorbente lodos cerámicos. Se evaluaron con los inputs pH de la solución de prueba, concentración inicial del Cu(II), tiempo de contacto con el

sorbente y la temperatura sometida al sistema. (Kashif Uddin, Ali Khan Rao, & Chandra Mouli, 2018)

Para la remoción de Zinc (II) en aguas utilizando como sorbente, se aplicó la RNA con evolución diferencial meta-heurística para el análisis predictivo y la superficie de respuesta para la optimización de los inputs: Concentración inicial de zinc, pH de la solución, tiempo de agitación, dosis de carbón activado, temperatura aplicada a la solución. (Rao Karri & Sahu, 2018).

Se utilizó como sorbente Cáscaras de caparazón de caracol en polvo soportado en carbonato de calcio para la remoción de cadmio (II) en aguas, utilizando como herramienta predictora a la RNA, teniendo como variables de entrada la concentración inicial de Cd(II), pH de la solución de interés, cantidad de sorbente y tiempo de contacto. (Gautam & Asif, 2018)

Utilizando un reactor Biochar, se estudió la remoción de iones cadmio, níquel, zinc y cobre en aguas. Como herramienta de predicción se utilizaron 4 tipos de RNA: (i) propagación hacia adelante con retroalimentación, (ii) propagación hacia adelante con retroalimentación con retardo de tiempo distribuido, (iii) RNA con cascada, (iv) RNA de Elman. (Moreno-Pérez, y otros, 2018).

Teniendo como interés las hojas de Ficus Carica como sorbente, se estudió la remoción de calcio (II), Cr(III), Co(II), Cu(II), Cd(II), K(I), Mg(II), Na(I), Ni(II), Pb(II), Zn(II), Fe(II) en aguas, utilizando como modelos de predicción: RNA, también la regresión lineal múltiple y la relación propiedad/estructura (QSPR). (Iqbal & Akbar, 2017).

Utilizando el aceite de pongamia (*Pongamia pinnata*) como sorbente de Zn(II) en aguas, se aplicó la RNA como herramienta predictora y la superficie de respuesta como solución óptima de la remoción de zinc, teniendo como variables de entrada el pH de la solución de interés, concentración inicial de Zn(II), temperatura, y dosis de sorbente. (Shanmugarprakash, Venkatachalam, Rajendran, & Pugazhendhi, 2018)

Utilizando materiales estructurados de metal poroso-orgánico, se probó la capacidad de remoción de metales en aguas, cuando dicha agua, en estado de vapor, tiene contacto con el sorbato. Para las herramientas de predicción se utilizó RNA y ANFIS. (Mahdi Niknam, Morteza, & Maryam, 2018)

Utilizando el flósculo de banano como biosorbente se evaluó la eliminación de los iones Cu(II) en aguas por medio de experimentación por lotes, evaluando los efectos de agitación, tamaño de partícula, pH, temperatura, y concentración inicial de Cu(II), teniendo como herramienta de predicción la RNA y de optimización el algoritmo genético. (Sutherland, Marcano, & Chitto, 2018).

Se citan además dos trabajos donde se aplica las RNA como herramienta predictora, y teniendo como sorbente el quitosano:

Considerando la espinela de ferrita soportado en quitosano entrecruzado con glutaraldehído, se estudió la remoción del colorante *Acid Orange 7* (AO7) en soluciones acuosas, con el modelo de predicción de la RNA y para la optimización los algoritmos genéticos, teniendo como inputs el pH, la concentración inicial del AO7, temperatura, y dosis del sorbente. (Cojocar, Samoila, & Pascariu, 2018)

Obteniendo el sorbente: nanodióxido/quitosano (CS-NDIO), se estudió la remoción del cristal violeta (CV) en aguas, a partir de un diseño de experimentos compuesto central y teniendo como herramienta de predicción a la RNA y como herramienta de optimización a los algoritmos genéticos, considerando como variables de entrada el pH, concentración inicial del CV, temperatura y dosis de sorbente. (Ghanavati Nasab, y otros, 2018)

OBJETIVOS

OBJETIVOS GENERALES

Desarrollar una red neuronal artificial que proporcione el soporte predictivo para la descontaminación de cobre (II) en aguas utilizando biopolímeros, teniendo como parámetros característicos del proceso químico al pH, masa de perlas de quitosano entrecruzadas con glutaraldehído, concentración de glutaraldehído utilizada para el entrecruzamiento, tiempo de agitación, frecuencia de agitación, y concentración inicial de cobre.

OBJETIVOS ESPECÍFICOS

1. Preparar las perlas de quitosano entrecruzadas con glutaraldehído para la remoción de Cu en soluciones acuosas.
2. Realizar los experimentos de remoción del ion Cu (II) utilizando las perlas de quitosano entrecruzadas con glutaraldehído, siguiendo la metodología del diseño experimental factorial con puntos centrales, considerando como factores experimentales: pH, masa de perlas de quitosano entrecruzadas con glutaraldehído, concentración de glutaraldehído utilizada para el entrecruzamiento, tiempo de agitación, frecuencia de agitación, y concentración inicial del cobre.
3. Diseñar una red neuronal artificial del tipo perceptrón para poder obtener el soporte predictivo sobre la información del porcentaje de remoción del contaminante.
4. Realizar el análisis estadístico de la red neuronal artificial diseñada, con el fin de determinar el desempeño de la misma como soporte predictivo.

MARCO TEÓRICO

1. QUITOSANO¹

1.1 Breve reseña histórica del Quitosano:

El quitosano es considerado un polímero heterogéneo de unidades D-glucosamina y N-acetil-D-glucosamina. El quitosano es un material altamente variable en sus características, entre ellos el porcentaje de desacetilación, masa molecular y cristalinidad; observándose discrepancias entre los diferentes estudios relacionados con la sorción de metales.

Estas características también pueden ser explicadas en el origen del material, ya que la quitina puede ser extraída de diferentes materiales biológicos, las cuales darán origen a tres tipos de quitina, y son definidas en relación con la orientación de sus cadenas: un arreglo antiparalelo se denomina α -quitina, forma más estable y abundante (generalmente se encuentra en la cutícula de artrópodos y exoesqueletos de cangrejos y camarones), un arreglo paralelo se denomina β -quitina, y una mezcla de arreglos, dos paralelos y uno antiparalelo se denomina γ -quitina (ambas formas obtenidas a partir del esqueleto del calamar). Estas tres formas generalmente diferenciadas en la dureza, siguiendo el orden: α -quitina, β -quitina, γ -quitina. Otra razón que puede afectar las características del material son las condiciones experimentales empleadas en la obtención de la quitina y quitosano. Esto básicamente puede afectar el porcentaje de desacetilación, la cristalinidad y la masa molecular del quitosano (Skjak-Braek, Anthonsen, & Sandford, 1989).

La desacetilación completa de la quitina produce un material totalmente soluble en medio ácido conocido como quitano (**Gráfico 3**); sin embargo, cuando la desacetilación del material de partida es incompleta se crea una mezcla de cadenas que tienen distintas proporciones de unidades $\beta(1-4)$ -2-acetamida-2-desoxi-D-glucosa y $\beta(1-4)$ -2-amino-2-desoxi-D-glucosa, cuya relación depende de las condiciones de reacción y que obviamente, genera materiales

¹ (Díaz Sánchez, 2011)

con distintas propiedades denominados quitosanos. La diferencia en las propiedades de estos materiales puede llegar a ser notable, como por ejemplo la distinta solubilidad en medio acuoso que pueden llegar a tener. Las estructuras químicas del quitano y el quitosano se muestran a continuación, en el **Gráfico 1**

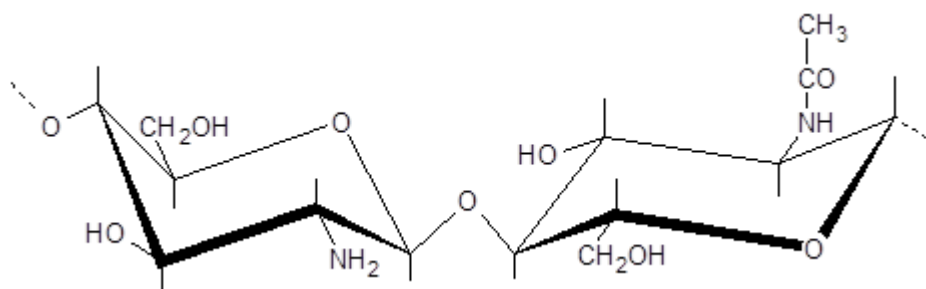


Gráfico 1: Unidad repetitiva del quitosano.

En el **Gráfico 2** se presenta el mecanismo de reacción para la obtención del quitosano.

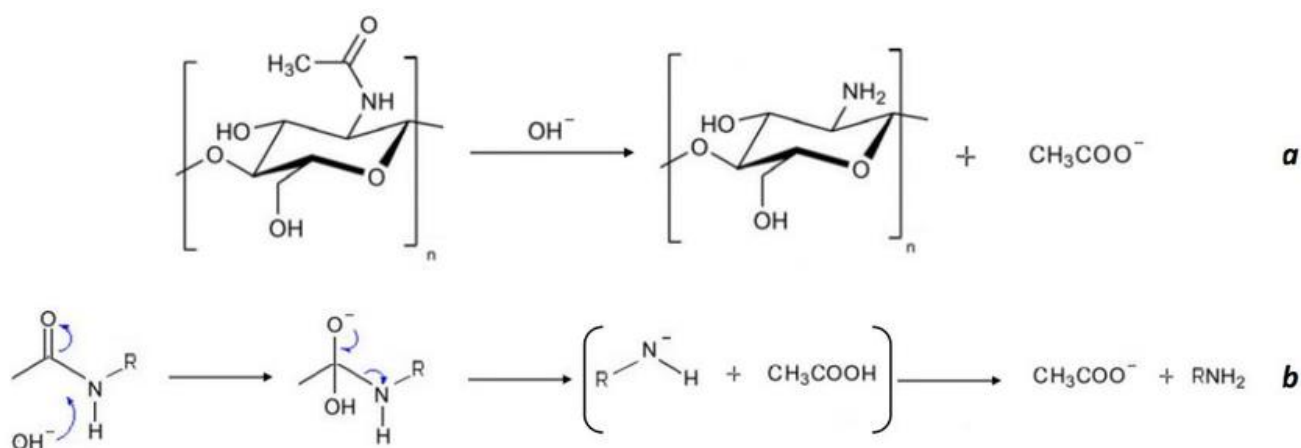


Gráfico 2: Mecanismo de reacción de la desacetilación de la quitina para obtener quitosano. a) Reacción generalizada, b) Detalle del mecanismo. (Marinela, 2014)

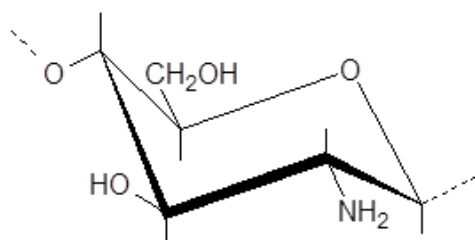


Gráfico 3: Unidad repetida de quitano

La presencia de grupos aminos en la cadena polimérica ha hecho del quitosano uno de los materiales más versátiles que se estudian desde hace ya algún tiempo, por la posibilidad de realizar una amplia variedad de modificaciones, tales como las reacciones de anclaje de enzimas, reacciones de injerto, obtención de películas entrecruzadas, etc., de las cuales se obtienen materiales con propiedades adecuadas para aplicaciones inmediatas y futuras en biotecnología, biomedicina, agricultura, etc.

Es importante recalcar la febril actividad que se desarrolla en torno a estos materiales, lo que se ha traducido en la aparición de Congresos, Simposios, Conferencias, etc., dedicados exclusivamente a ellos. Su importancia adquirió matiz internacional como materiales poliméricos de primer orden, especialmente por su biodegradabilidad.

1.2 Química del quitosano:

1.2.1 Reacciones de entrecruzamiento:

Los hidrogeles químicos son redes tridimensionales formadas por la unión covalente de todas las cadenas de polímeros existentes en un sistema determinado, hasta formar, teóricamente, una sola macromolécula. El quitosano es uno de los biomateriales más extensamente estudiados en este sentido. Existen varios agentes entrecruzantes que han sido estudiados para tal fin, como por ejemplo la epíclorhidrina y el etilenglicol glicidil éter. Pero quizás la reacción de entrecruzamiento más frecuentemente utilizada para este fin es la reacción con dialdehídos para formar uniones químicas tipo base de Schiff (Crescenzi, Imbriaco, Lárez, Dentini, & Ciferri, 1995). Las bases de Schiff (*II*) son especies generadas a partir de reacciones reversibles; sin embargo, estas especies pueden ser convertidas en materiales más estables (*III*) mediante una reacción de hidrogenación con reactivos comunes como el cianoborohidruro de sodio, de acuerdo con el siguiente esquema:

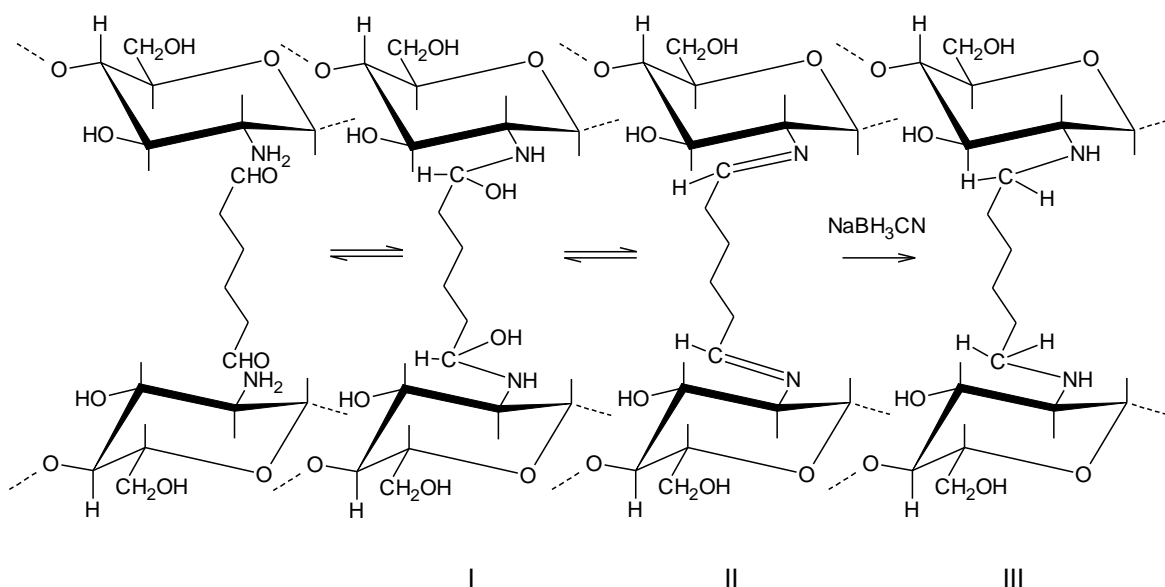


Gráfico 4: Reacciones de entrecruzamiento para producir hidrogeles de quitosano usando un dialdehído como agente entrecruzante. (Díaz Sánchez, 2011)

Obviamente un polímero con grupos aldehídos es candidato obligado para la preparación de este tipo de hidrogeles, especialmente si dicho material proviene de fuentes naturales como el biopolímero denominado scleroglucano (*Polytran, Biopolímero CS*), el cual es secretado por hongos del género *Sclerotium*, entre otros. Este material ha sido oxidado exitosamente a un derivado conocido como sclerodialdehído mediante la conocida reacción de grupos hidroxilo vecinales con periodato de sodio. Posteriormente, a través de una reacción similar a la mostrada en el esquema anterior, se ha logrado producir un hidrogel que ha sido estudiado en sus propiedades de hinchamiento en función del pH del medio, absorción de compuestos aniónicos como el ácido salicílico, cationes metálicos y también su interacción con surfactantes, donde fue posible observar el llamado colapso del sistema (Lárez, Crescenzi, Dentini, & Ciferri, 1995).

La formación de redes tridimensionales con quitosano también se ha realizado con diversos agentes entrecruzantes macromoleculares, como por ejemplo mediante reacciones de injerto del poliácido acrílico a través del grupo amina.

2. EL COBRE

2.1 Química del Cobre

El cobre (del latín *cuprum*, y éste del griego *kypros*), cuyo símbolo es *Cu*, es el elemento químico de número atómico 29. Se trata de un metal de transición de color cobrizo (rojizo) y brillo metálico que, junto con la plata y el oro, forma parte de la llamada familia del cobre, se caracteriza por ser uno de los mejores conductores de electricidad (el segundo después de la plata). Gracias a su alta conductividad eléctrica, ductilidad y maleabilidad, se ha convertido en el material más utilizado para fabricar cables eléctricos y otros componentes eléctricos y electrónicos.

Tabla 1: Propiedades químicas del Cu en el ambiente

Nombre	Cobre
Número atómico	29
Valencia	1.2
Estado de oxidación	2+
Electronegatividad	1.9
Radio covalente (Å)	1.38
Radio iónico (Å)	0.69
Radio atómico (Å)	1.28
Configuración electrónica	[Ar]3d ¹⁰ 4s ¹
Primero potencial de ionización (eV)	7.77
masa atómica (g/mol)	63.546 ± 0.003
Densidad (g/mL)	8.96
Punto de fusión (°C)	1083
Punto de ebullición (°C)	2595

2.2 El cobre como contaminante

El cobre es un nutriente esencial y, al mismo tiempo, un contaminante del agua de consumo. Tiene muchos usos comerciales: se utiliza para fabricar tuberías, válvulas y accesorios de fontanería, así como en aleaciones y revestimientos. En ocasiones se añade sulfato de cobre pentahidratado a las aguas superficiales para el control de algas. Las concentraciones de

cobre en el agua de consumo varían mucho, y la fuente principal más frecuente es la corrosión de tuberías de cobre interiores. Las concentraciones suelen ser bajas en muestras de agua corriente o que se ha dejado correr prolongadamente, mientras que en muestras de agua retenida o que se ha dejado correr poco tiempo son más variables y suelen ser considerablemente más altas (con frecuencia >1 mg/L). La concentración de cobre en el agua tratada suele aumentar durante su distribución, sobre todo en sistemas con pH ácido o en aguas con concentración alta de carbonato, con pH alcalino. Las fuentes principales de exposición al cobre en los países desarrollados son los alimentos y el agua. El consumo de agua retenida o que se ha dejado correr poco tiempo de sistemas de distribución con tuberías o accesorios de cobre puede hacer aumentar considerablemente la exposición diaria total al cobre, especialmente en lactantes alimentados con leche maternizada en polvo reconstituida con agua de grifo. La **Tabla 2** recoge la información general de la OMS.

Tabla 2: Datos de referencia del Cu como contaminante. (OMS, *Cooper in drinking-water*, 2003)

Valor de referencia	2 mg/L
Presencia	Las concentraciones en el agua de consumo varían de ≤ 0.005 hasta > 30 mg/L, principalmente como resultado de la corrosión interior de tuberías de cobre.
Método de determinación del valor de referencia	Para proteger de los efectos gastrointestinales agudos del cobre y proporcionar un margen de seguridad adecuado a las poblaciones con una homeostasis normal del cobre.
Límite de detección (referenciales, en promedio)	0.02 – 0.1 $\mu\text{g/L}$ mediante ICP-MS; 0.3 $\mu\text{g/L}$ mediante ICP-AES; 0.5 $\mu\text{g/L}$ mediante FAAS.
Concentración alcanzable mediante tratamiento	Los tratamientos convencionales no eliminan el cobre. No obstante, el cobre no es un contaminante habitual del agua bruta.
Observaciones adicionales	<ul style="list-style-type: none"> El valor de referencia debería permitir a las personas adultas con una homeostasis normal del cobre beber de 2 a 3 litros de agua al día y consumir cobre en complementos alimenticios y en los

	<p>alimentos sin exceder la ingesta máxima tolerable de 10 mg/día ni provocar una respuesta gastrointestinal adversa.</p> <ul style="list-style-type: none">• Cuando la concentración de cobre es superior a 1 mg/L, el agua mancha la ropa lavada y los aparatos sanitarios. En concentraciones superiores a 2,5 mg/l, el cobre confiere un sabor amargo no deseado al agua; en concentraciones superiores afecta también a su color.• En la mayoría de los casos en los que se utilizan tuberías de cobre como material de fontanería, la concentración de cobre será inferior al valor de referencia. No obstante, en determinadas circunstancias, como en el caso de las aguas muy ácidas o corrosivas, se generarán concentraciones de cobre mucho más altas, y la utilización de tuberías de cobre puede no ser apropiada.
--	---

2.3 Reseña toxicológica

El IPCS concluyó que el límite máximo aceptable de ingesta por vía oral para adultos resulta dudoso, pero es probable que sea del orden de varios miligramos al día (más de 2 o 3, pero no muchos). Esta evaluación se basó únicamente en estudios sobre los efectos gastrointestinales del agua de consumo contaminada con cobre. Se consideró que los datos disponibles de toxicidad en animales no eran de utilidad para establecer el límite máximo aceptable de ingesta por vía oral, debido a la incertidumbre sobre el modelo adecuado para el ser humano, aunque se tuvieron en cuenta para determinar un modo de acción de la respuesta. Los datos sobre los efectos gastrointestinales del cobre deben emplearse con precaución, ya que la concentración del cobre ingerido influye más en los efectos observados que la masa total o dosis ingerida durante 24 horas. En estudios recientes se ha definido el umbral de concentración de cobre en el agua de consumo que produce efectos sobre el aparato digestivo, pero todavía hay ciertas dudas respecto a los efectos del cobre a largo

plazo en poblaciones sensibles, como los portadores del gen de la enfermedad de Wilson o los afectados por otros trastornos metabólicos de la homeostasis del cobre. (IPCS, 1998)

2.4 Valor máximo permisible de cobre en el Perú (Ministerio del Ambiente, 2017)

En el Perú, el Ministerio del Ambiente (MINAM), a través del Estándar de Calidad Ambiental (ECA – 07/06/2018) de aguas regula la contaminación del cuerpo acuoso a través de límites máximos permisibles (LMP). En cuestión del cobre en aguas, las ECA's aguas del MINAM indican los siguientes límites (**Tabla 3**):

Tabla 3: Límite Máximo Permisible de Cobre según legislación peruana (MINAM) - ECA aguas 2017

Aplicabilidad	LMP	Unidad
Categoría 1: Poblacional y recreacional		
Subcategoría A: Aguas destinadas a la producción de agua potable		
A1: Aguas que pueden ser potabilizadas con desinfección	2.0	mg/L
A2: Aguas que pueden ser potabilizadas con tratamiento convencional	2.0	mg/L
A3: Aguas que pueden ser potabilizadas con tratamiento avanzado	2.0	mg/L
Subcategoría B: Aguas superficiales destinadas para recreación		
B1: Contacto primario	2.0	mg/L
B2: Contacto secundario	No aplica	No aplica
Categoría 2: Extracción, cultivo y otras actividades marino costeras y continentales		
C1: Extracción y cultivo de moluscos, equinodermos y tunicados en aguas marino costeras	0.0031	mg/L
C2: Extracción y cultivo de otras especies hidrobiológicas en aguas marino costeras	0.05	mg/L
C3: Actividades marino portuarias, industriales o de saneamiento en agua marino costeras	0.05	mg/L
C4: Extracción y cultivo de especies hidrobiológicas en lagos o lagunas	0.2	mg/L
Categoría 3: Riego de vegetales y bebida de animales		

D1: Riego de vegetales - agua para riego no restringido	0.2	mg/L
D1: Riego de vegetales - agua para riego restringido	0.2	mg/L
D2: Bebida de animales	0.5	mg/L
Categoría 4: Conservación de ambiente acuático		
E1: Laguna y lagos	0.1	mg/L
E2: Ríos - Costa y sierra	0.1	mg/L
E2: Ríos - Selva	0.1	mg/L
E3: Ecosistemas costeros y marinos - Estuarios	0.05	mg/L
E3: Ecosistemas costeros y marinos - Marinos	0.05	mg/L

3. BIOSORCIÓN

El fenómeno de biosorción (Volesky, 1986) ha demostrado actualmente ser una alternativa para el tratamiento de diferentes efluentes industriales con respecto a otros métodos fisicoquímicos (precipitación, floculación-coagulación, intercambio iónico, separación por membranas, etc.). En el proceso de biosorción se utilizan materiales de origen biológico, tales como: Algas, hongos, caparazón de artrópodos, bacterias, restos de vegetales, etc., los cuales se encuentran en gran abundancia y son fácilmente transformables a biosorbentes. Los iones de los metales pesados se unen a los centros activos de adsorción del material biológico mediante la formación de complejos, quelatos, intercambio iónico, microprecipitación en la parte interna del material, etc.

Por otro lado, es necesario señalar que en la extracción de iones de metales pesados de efluentes industriales también se está utilizando la biomasa viva, pero esta técnica tiene serias limitaciones, por cuanto los iones al acumularse en la pared celular en la parte interna de la célula impiden el normal desarrollo de las plantas y microorganismos que en muchos casos pueden conducir a alteraciones genéticas. Este problema no se presenta en la biomasa

muerta ya que esta se puede modificar física y químicamente para mejorar su capacidad de atrapar a los iones de metales pesados que son contaminantes del medio ambiente.

3.1 Biosorción con quitosano

La contaminación de las aguas es uno de los aspectos más preocupantes de la degradación de los medios naturales por parte de la civilización contemporánea (Jean & Holl, 2003). Los estudios sobre aguas residuales confirman que la contaminación tiene, generalmente, un origen químico. Los principales agentes contaminantes son pesticidas, hidrocarburos y metales pesados. Los metales pueden provenir de efluentes urbanos (alimentos, productos farmacéuticos, productos de limpieza, etc.) y de fuentes industriales (industrias papeleras, curtiembres, pinturas y pigmentos, etc.).

La recuperación de metales de efluentes acuosos puede llevarse a cabo por diferentes métodos (Zhan & Banks, 2005). Además de los tratamientos mecánicos de las aguas residuales (sedimentación) o de los biológicos (lodos activados), se utilizan algunos tratamientos químicos para la eliminación de estos metales pesados. Los procesos más comunes implican la precipitación mediante hidróxidos o sulfuros, la oxidación-reducción, el intercambio iónico, la separación sólido-líquido mediante decantación-flotación y la separación mediante membranas. Pero el mayor inconveniente de estos tratamientos es la formación de lodos que tienen que estar sujetos a restricciones o la baja eficiencia del proceso en el caso de las membranas. De aquí surge la necesidad de buscar procesos alternativos más económicos basados, por ejemplo, en la utilización de polímeros naturales o sintéticos. La adsorción en carbón activo ha sido ampliamente estudiada. Se han considerado también otros adsorbentes con el fin de encontrar materiales más eficientes y menos costosos. Así, se está estudiando la bioadsorción en organismos vivos (bacterias,

hongos, algas) o por compuestos extraídos de estos organismos (Tsezos, 2001). Estos materiales son abundantes y biodegradables. Entre estos últimos se encuentra el quitosano.

El quitosano, poli $\beta(1-4)$ -2-amino-2-desoxi-D-glucosa, se prepara por desacetilación de los grupos acetamida de la quitina, que es un polímero natural extraído de los caparazones de crustáceos, tales como cangrejos, insectos y camarones. La capacidad del quitosano para formar complejos con distintos iones metálicos presenta un gran interés para los investigadores (*Benaissa & Benguella, 2002*). La comparación entre los diferentes estudios realizados es complicada debido a la gran variabilidad de condiciones experimentales empleadas. Este polímero se caracteriza por poseer un elevado número de grupos aminos libres que son muy reactivos para la quelación de cationes metálicos a pH más o menos neutros (*Alsarra & al, 2002*).

Modificaciones físicas y químicas se han desarrollado para mejorar su calidad, capacidad y selectividad adsorbente del quitosano. Por ejemplo, la cuaternización del grupo amino y el entrecruzamiento con glutaraldehído para mejorar su resistencia y estabilidad en medios ácidos.

3.2 Propiedades de acomplejación

La capacidad del quitosano para formar complejos con iones metálicos es una de las propiedades más destacables de este polímero. Los grupos amino, mediante enlaces dativos, son los principales sitios de interacción con los metales, si bien los grupos hidroxilo también pueden participar en este vínculo. La capacidad de interacción entre quitosano y los metales varía con el grado de desacetilación, el número de grupos amino y el peso molecular del polímero. Por otro lado, las características de la disolución en la que se realiza la acomplejación (pH, proporción entre ligando y metal, temperatura) también influyen determinadamente en la acomplejación.

No se ha llegado a elucidar completamente el mecanismo mediante el cual el ion metálico se une al polímero. La bibliografía propone distintos mecanismos como pueden ser la quelación, la absorción o el intercambio iónico.

Existen dos modelos principales para describir la unión entre el quitosano y los metales. El primero es el denominado modelo colgante, en el que el metal está unido a un único grupo amino de la cadena polimérica. En el otro modelo se hace referencia a una unión tipo puente en la que el metal se coordina con varios átomos de nitrógeno ya sea de una o de varias cadenas del polímero. Las representaciones de ambos modelos se observan en el **Gráfico 5**.

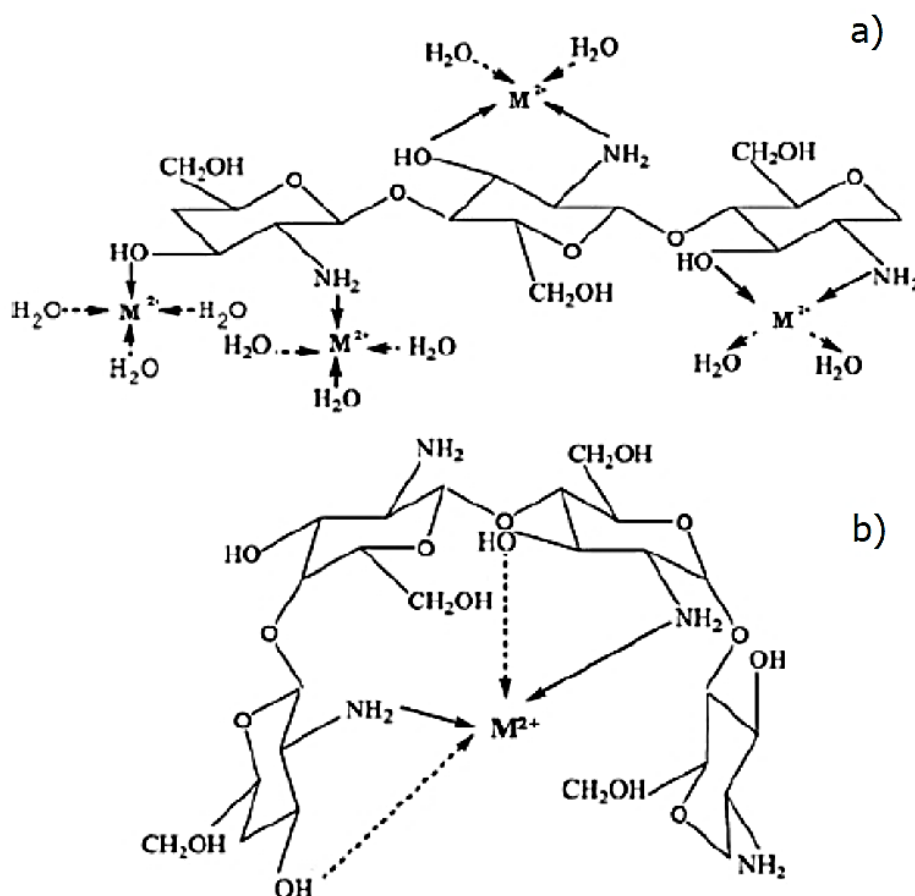


Gráfico 5: Modelos de acomplejación de metales tipo a) colgante y b) puente. (Lasheras Zubiate, 2011)

4. REDES NEURONALES ARTIFICIALES²

4.1 Introducción

Las Redes Neuronales Artificiales (RNA) surgieron originalmente como una simulación abstracta de los sistemas nerviosos biológicos, constituidos por un conjunto de unidades llamadas neuronas conectadas unas con otras. El primer modelo de red neuronal fue propuesto en 1943 por McCulloch y Pitts en términos de un modelo computacional de actividad nerviosa. Este modelo era un modelo binario, donde cada neurona tenía un escalón o umbral prefijado, y sirvió de base para los modelos posteriores.

4.2 Definición

Una *Red Neuronal Artificial (RNA)* es un sistema para el tratamiento de la información, cuya unidad básica de procesamiento está inspirada en la célula fundamental del sistema nervioso humano: la neurona.

Las neuronas son un componente relativamente simple pero conectadas de a miles forman un poderoso sistema.

4.2.1 Descripción de las RNA

- Unidades de procesamiento que intercambian datos o información.
- Se utilizan para reconocer patrones, incluyendo imágenes, manuscritos, tendencias financieras, etc.
- Tienen la capacidad de aprender y mejorar su funcionamiento.

4.2.2 Comparación con otra técnica de Inteligencia Artificial

A continuación, se muestra un esquema de comparación entre sistemas basados en inteligencia, tabulando inmediatamente las similitudes y diferencias entre redes neuronales

² (Russell & Norving, Inteligencia Artificial un enfoque moderno, 2004)

artificiales y los sistemas expertos. La **Tabla 4** presenta una comparación entre las RNA y los sistemas expertos.

Tabla 4: Comparación entre redes neuronales artificiales y sistemas expertos (*Slater, Hazen, & Sakthivel, 1993*).

SEMEJANZAS	DIFERENCIAS	
	RNA	Sistemas Expertos
Los dos manejan sistemas de análisis.	Tienen capacidades para que los usuarios los desarrollen, y pueden hacerlo incluso aquellos con poca formación.	Tienen capacidades para que los usuarios los desarrollen, pero es preferible que lo hagan profesionales cualificados dadas las complicaciones que entraña la adquisición de conocimientos.
Los dos sistemas son digitales e intangibles.	Pueden desarrollarse en poco tiempo. Los expertos únicamente tienen que identificar los datos a introducir, los resultados, y una amplia gama de muestras.	Su desarrollo lleva mucho más tiempo. Hay que tener expertos disponibles y dispuestos a articular el proceso de resolución de problemas.
En los dos sistemas se encargan de representar u obtener el comportamiento y conocimiento humano	No hay que identificar las reglas. Muy adecuadas para decisiones tomadas de forma intuitiva.	Las reglas deben estar claramente identificadas. Difíciles de desarrollar en aquellas decisiones que se toman de forma intuitiva.
En los sistemas se utilizan variables / reglas y por medio de ellos se mejoran las tareas que se terminan convirtiendo en Software.	Muy adecuadas para reconocimientos de patrones, pero necesitan una amplia gama de datos de nuestra. Muy tolerantes a los fallos. Muy adáptales a entornos de problemas cambiantes.	Torpes en el reconocimiento de patrones y el análisis de datos, como los pronósticos No son tolerantes a los fallos. Los cambios en el entorno de problemas garantizan el mantenimiento.

Se puede intentar, con las redes de neuronas artificiales, si la aplicación no se ajusta a uno de los esquemas de representación del sistema experto.	La aplicación debe ajustarse a uno de los esquemas de representación de conocimientos.
Las redes de neuronas artificiales superan a los expertos humanos en determinadas aplicaciones, como los pronósticos.	El rendimiento del experto humano que ayudó a crear el sistema experto establece un límite en el rendimiento teórico de este último.
No tienen sistema de explicación, y actúan como caja negra.	Tienen sistemas de explicación para justificar por qué y cómo se tomó la decisión. Necesarios cuando hay que explicar la decisión para inspirar confianza a los usuarios. Recomendados cuando el proceso de resolución de problemas es bien conocido.
Útiles para decisiones repentinas.	Útiles cuando se toman una serie de decisiones en forma de árbol y cuando en esos casos es necesaria la interacción de un usuario.
Útiles cuando es necesario emular funciones humanas de bajo nivel, como el reconocimiento de patrones.	Útiles cuando es necesario emular funciones humanas de alto nivel, como el razonamiento y la deducción.
Las redes de neuronas artificiales en ciertos casos son útiles para validar la corrección del desarrollo de un sistema experto.	Los sistemas expertos no son útiles para dar validez a la corrección desarrollo de un sistema de redes neuronales artificiales.

4.3 Idea de red neuronal artificial

4.3.1 Redes Neuronales Artificiales (RNA)

Las RNA aplicadas están, en general, inspiradas en las redes neuronales biológicas (**Gráfico 6**), aunque poseen otras funcionalidades y estructuras de conexión distintas a las vistas desde la perspectiva biológica. Las características principales de las RNA son las siguientes:

- 1. Auto-Organización y Adaptabilidad:** Utilizan algoritmos de aprendizaje adaptativo y auto-organización, por lo que ofrecen mejores posibilidades de procesamiento robusto y adaptativo.
- 2. Procesado no Lineal:** Aumenta la capacidad de la red para aproximar funciones, clasificar patrones y aumenta su inmunidad frente al ruido.
- 3. Procesado Paralelo:** Normalmente se usa un gran número de nodos de procesamiento, con alto nivel de interconectividad.

El elemento básico de computación (modelo de neurona) es un nodo o unidad (**Gráfico 7**). Recibe un input desde otras unidades o de una fuente externa de datos. Cada input tiene un peso asociado w , que se va modificando en el llamado proceso de aprendizaje. Cada unidad aplica una función dada f de la suma de los inputs ponderadas mediante los pesos.

$$y_i = \sum_j w_{ij} y_j \quad \text{Ecuación 1}$$

El resultado puede servir como output de otras unidades.

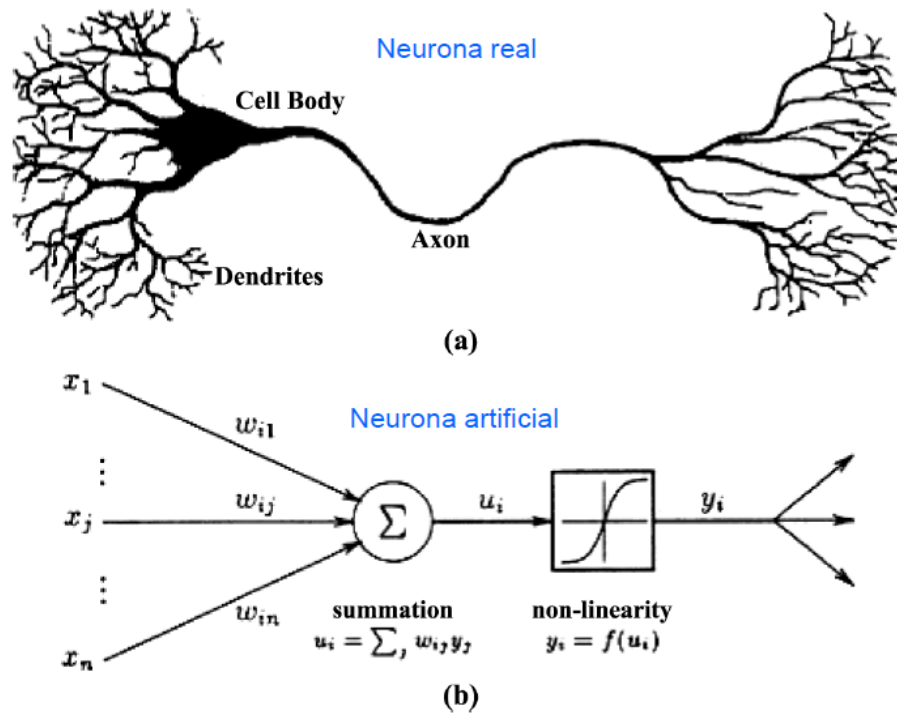


Gráfico 6: Similitud operativa entre una neurona biológica y una neurona artificial.

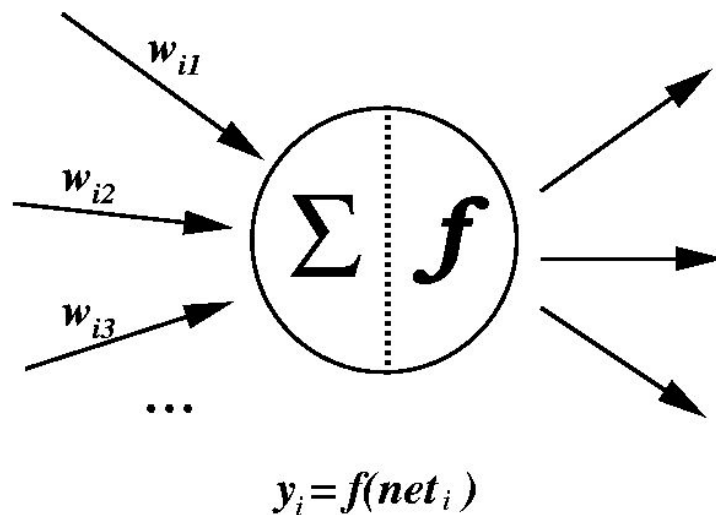


Gráfico 7: Esquema funcional del núcleo de la neurona artificial.

Las características de las RNA juegan un importante papel, por ejemplo, en el procesamiento de señales e imágenes. Se usan arquitecturas que comprenden elementos de procesamiento adaptativo paralelo, combinados con estructuras de interconexiones jerárquicas.

Hay dos fases en la modelización con redes neuronales:

- **Fase de entrenamiento:** se usa un conjunto de datos o patrones de entrenamiento para determinar los pesos (parámetros) que definen el modelo de red neuronal. Se calculan de manera iterativa, de acuerdo con los valores de los valores de entrenamiento, con el objeto de minimizar el error cometido entre la salida obtenida por la red neuronal y la salida deseada.
- **Fase de Prueba:** en la fase anterior, el modelo puede que se ajuste demasiado a las particularidades presentes en los patrones de entrenamiento, perdiendo su habilidad de generalizar su aprendizaje a casos nuevos (sobreajuste).

Para evitar el problema del sobreajuste, es aconsejable *utilizar un segundo grupo de datos diferentes a los de entrenamiento*, el grupo de validación, que permita controlar el proceso de aprendizaje.

Normalmente, los pesos óptimos se obtienen optimizando (minimizando) alguna función de *energía* (impulso).

Por ejemplo, un criterio muy utilizado en el llamado entrenamiento supervisado, que consiste en *minimizar el error cuadrático medio* entre el valor de salida y el valor real esperado.

Tabla 5: Algoritmos de aprendizaje para tareas de análisis de datos.

Paradigma	Regla de aprendizaje	Arquitectura	Algoritmo de aprendizaje	tareas
Supervisado	Corrección del error	Perceptrón o perceptrón multicapa	Algoritmos de aprendizaje perceptrón, retropropagación del error, ADALINE, Madeline	Clasificación de patrones, aproximación de funciones, predicción, control
		Elman y Jordan recurrentes	Retropropagación del error	Síntesis de series temporales
	Boltzmann	Recurrente	Algoritmo de aprendizaje Boltzman	Clasificación de patrones

	Competitivo	Competitivo	LVQ	Categorización intra-clase, comprensión de datos
		Red ART	ARTMap	Clasificación de patrones, categorización intra-clase
No supervisado	Corrección del error	Red de Hopfield	Aprendizaje de memoria asociativa	Memoria asociativa
		Multicapa sin realimentación	Proyección de Sannon	Análisis de datos
	Competitiva	Competitiva	VQ	Categorización, comprensión de datos.
		SOM	Kohonen SOM	Categorización, análisis de datos
		Redes ART	ART1, ART2	Categorización
Por refuerzo	Hebbian	Multicapa sin realimentación	Análisis lineal de discriminante	Análisis de datos, clasificación de patrones
		Sin realimentación o competitiva	Análisis de componentes principales	Análisis de datos, comprensión de datos

4.3.2 Redes Neuronales Supervisadas y No Supervisadas

Las redes neuronales se clasifican comúnmente en términos de sus correspondientes algoritmos o métodos de entrenamiento: redes de pesos fijos, redes no supervisadas, y redes de entrenamiento supervisado. Para las redes de pesos fijos no existe ningún tipo de entrenamiento.

a. Reglas de entrenamiento Supervisada

Las redes neuronales de entrenamiento supervisado son las más populares. Los datos para el entrenamiento están constituidos por varios pares de patrones de entrenamiento de entrada y de salida. El hecho de conocer la salida implica que el entrenamiento se beneficia de la supervisión de un maestro. Dado un nuevo patrón de entrenamiento, en la etapa $(m + 1)$ -ésima, los pesos se adaptan de la siguiente forma:

$$w_{ij}^{m+1} = w_{ij}^m + \Delta w_{ij}^m \quad \text{Ecuación 2}$$

Se puede ver un diagrama esquemático de un sistema de entrenamiento supervisado en el **Gráfico 8**:

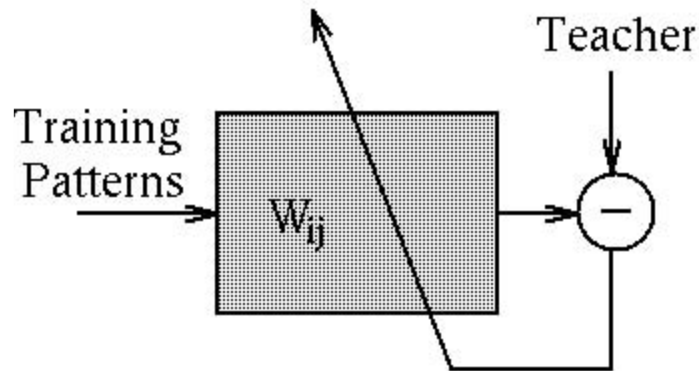


Gráfico 8: Esquema de sistema de entrenamiento supervisado.

b. Reglas de Entrenamiento No Supervisada

Para los modelos de entrenamiento no supervisado, el conjunto de datos de entrenamiento consiste sólo en los patrones de entrada. Por lo tanto, la red es entrenada sin la supervisión de un maestro. La red aprende a adaptarse basada en las experiencias recogidas de los patrones de entrenamiento anteriores. El siguiente esquema es el típico de un sistema No Supervisado se ilustra en el **Gráfico 9**:

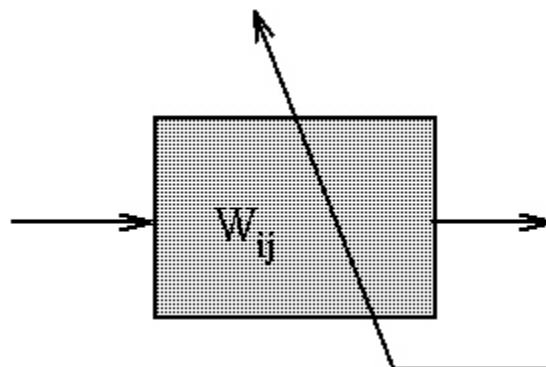


Gráfico 9: Esquema de sistema de entrenamiento no supervisado.

Por ejemplo, se aplican la Regla de Aprendizaje de Hebb y la Regla de Aprendizaje Competitivo. Así, en el primer caso se refuerza el peso que conecta dos nodos que se excitan simultáneamente.

En el aprendizaje competitivo, si un patrón nuevo pertenece a una clase ya presente, entonces la inclusión de este nuevo patrón a esta clase matiza la representación de la misma. Si el nuevo patrón no pertenece a ninguna de las clases anteriores, entonces la estructura y los pesos de la red neuronal serán ajustados para generar una nueva clase.

4.3.3 Funciones de Base y de Activación

Una red neuronal típica se puede caracterizar por la función base y la función de activación. Cada nodo (unidad de proceso), produce un valor y_j en su salida. Este valor se propaga a través de la red mediante conexiones unidireccionales hacia otros nodos de la red. Asociada a cada conexión hay un peso $\{w_{ij}\}$, que determina el efecto del nodo j -ésimo sobre el siguiente nodo i -ésimo.

Las entradas al nodo i -ésimo que provienen de otros nodos se suman junto con el valor umbral i (que hace el papel del término independiente de la regresión), y se aplica la función base f , obteniendo el valor u_i .

Finalmente se obtiene el valor y_i aplicando la función de activación sobre u_i .

a. Tipos de Función Base

La función base tiene dos formas típicas:

- **Función lineal de tipo hiperplano:** El valor de red es una combinación lineal de las entradas.

$$u_i(w, x) = \sum_{j=1}^n w_{ij}x_j \quad \text{Ecuación 3}$$

- **Función radial de tipo hiperesférico:** es una función base no lineal.

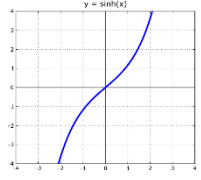
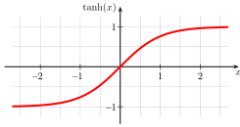
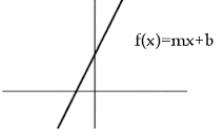
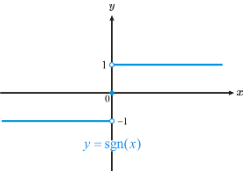
El valor obtenido representa la distancia a un determinado patrón de referencia.

$$u_i(w, x) = \sqrt{\sum_{j=1}^n (x_j - w_{ij})^2} \quad \text{Ecuación 4}$$

b. Función de Activación

El valor de red, expresado por la función de base, $u(w; x)$, se transforma mediante una función de activación no lineal. Las funciones de activación más comunes (**Tabla 6**) son la función logística y la función tangente hiperbólica:

Tabla 6: Funciones de activación más comúnmente usadas.

FUNCIÓN	DEFINICIÓN	Dom(F)	Ran(F)	Plot
Logística o seno hiperbólico	$f(x) = \text{senh}(x) = \frac{1}{1 + e^{-x}}$	\mathbb{R}	\mathbb{R}	
Tangente hiperbólica	$f(x) = \text{tanh}(x) = \frac{1 - e^{-x}}{1 + e^{-x}}$	\mathbb{R}	$[-1; 1]$	
Identidad	$f(x) = \text{lin}(x) = x$	\mathbb{R}	\mathbb{R}	
Signo	$f(x) = \text{sign}(x) = \begin{cases} 1; & x > 0 \\ 0; & x = 0 \\ -1; & x < 0 \end{cases}$	\mathbb{R}	$\{-1; 0; 1\}$	

4.3.4 Estructuras de conexión de atrás hacia delante

Una red neuronal se determina por las neuronas y la matriz de pesos.

Se pueden definir tres tipos de capas de neuronas:

- la capa de entrada,
- la capa oculta y
- la capa de salida.

Entre dos capas de neuronas existe una red de pesos de conexión, que puede ser de los siguientes tipos: hacia delante (feedforward), hacia atrás (feedback), lateral y de retardo (**Gráfico 10**).

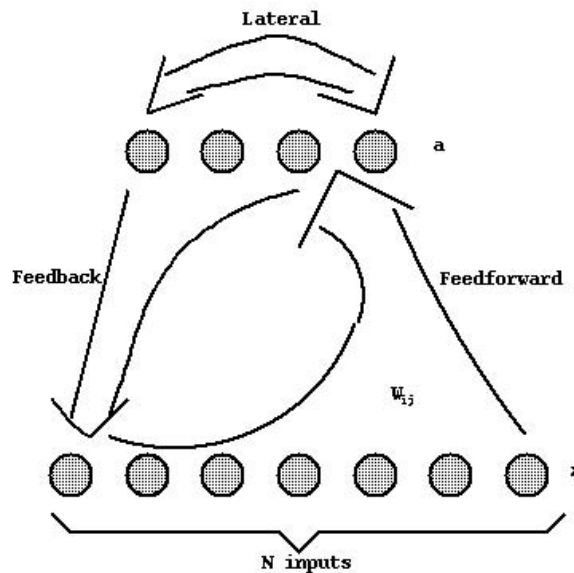


Gráfico 10: Tipos de redes de pesos de conexión entre dos capas de neuronas.

1. **Conexiones hacia delante:** los valores de las neuronas de una capa inferior son propagados hacia las neuronas de la capa superior por medio de las redes de conexiones hacia adelante.
2. **Conexiones hacia atrás:** estas conexiones llevan los valores de las neuronas de una capa superior a otras de la capa inferior.

3. **Conexiones laterales:** Un ejemplo típico de este tipo es el circuito “el ganador toma todo” (*winner-takes-all*), que cumple un papel importante en la elección del ganador: a la neurona de salida que da el valor más alto se le asigna el valor total (por ejemplo, 1), mientras que a todas las demás se le da un valor de 0.
4. **Conexiones con retardo:** los elementos de retardo se incorporan en las conexiones para implementar modelos dinámicos y temporales, es decir, modelos que precisan de *memoria*.

Las conexiones pueden ser total o parcialmente interconectadas, como se muestra en el **Gráfico 11**:

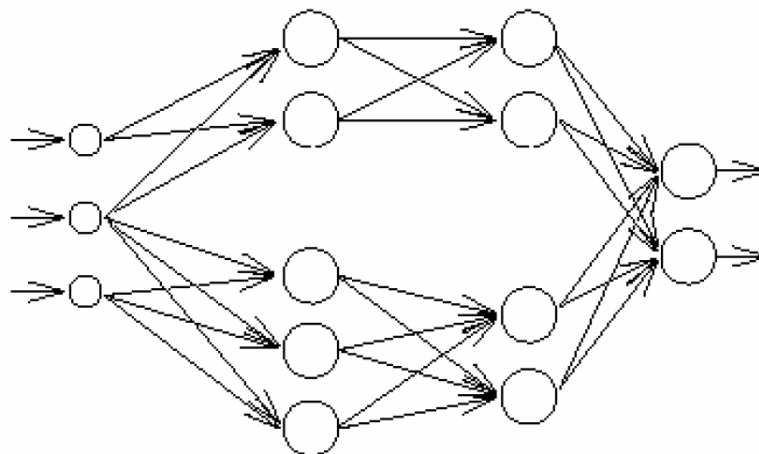


Gráfico 11: Red neuronal parcialmente interconectada.

También es posible que las redes sean de una capa con el modelo de pesos hacia atrás o bien el modelo multicapa hacia adelante. Es posible así mismo, el conectar varias redes de una sola capa para dar lugar a redes más grandes.

4.3.5 Tamaño de las Redes Neuronales

En una red multicapa de propagación hacia delante, puede haber una o más capas ocultas entre las capas de entrada y salida. El tamaño de las redes depende del número de capas y del número de neuronas ocultas por capa.

El número de unidades ocultas está directamente relacionado con las capacidades de la red. Para que el comportamiento de la red sea correcto, se tiene que determinar apropiadamente el número de neuronas de la capa oculta.

4.4 Métodos de regresión y redes neuronales

Los modelos de regresión estudian la relación entre una serie de variables, denominadas independientes x_i , y otras variables dependientes o respuesta que se denotan como y : El objetivo es predecir los valores de y en función de los valores de x_i :

El modelo básico se puede expresar como:

$$\eta = \sum_{i=0}^N \beta_i x_i \quad \text{Ecuación 5}$$

Donde:

$$E(y) = \mu \quad \text{Ecuación 6}$$

$$\mu = h(\eta) \quad \text{Ecuación 7}$$

En estas expresiones $h(\cdot)$ es la función que liga los componentes, β_i son los coeficientes, N es el número de variables independientes y β_0 es la pendiente.

Un modelo lineal se puede implementar como una red neuronal simple: tiene una unidad de sesgo, una unidad de input y una unidad de salida. El input se refiere a una variable x , mientras que el sesgo siempre es una constante igual a 1 (**Gráfico 12**). El output sería:

$$y_2 = y_1 w_{21} + 1,0 y_{20} \quad \text{Ecuación 8}$$

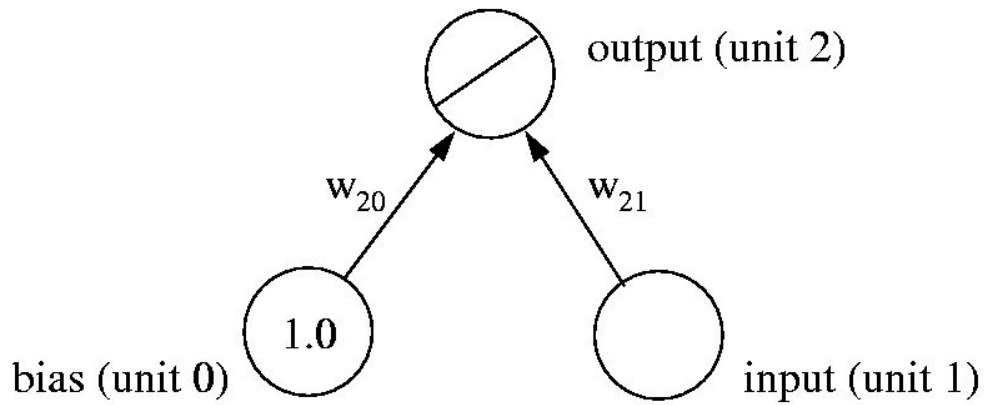


Gráfico 12: Red neuronal de un input (ejemplo).

En contrapartida, el **Gráfico 13** ilustra una RNA de más de un input:

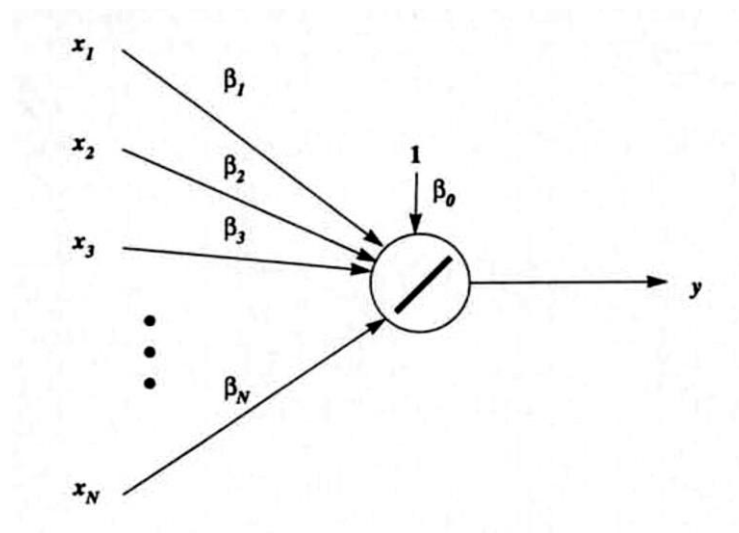


Gráfico 13: Red neuronal de varios inputs (ejemplo).

El modelo tiene tres componentes:

1. Un componente aleatorio de la variable respuesta y con media μ y varianza σ^2 :
2. Un componente que relaciona las variables independientes x_i con una función lineal

$$\eta = \sum_{i=0}^N \beta_i x_i.$$
3. Una función que relaciona la media con el predictor lineal $\eta = h^{-1}(\mu)$.

El modelo lineal generalizado se reduce al modelo de regresión lineal múltiple si se considera que el componente aleatorio se distribuye como una normal y la función $h(\cdot)$ se asume que es la identidad:

$$y_p = \beta_0 + \sum_{i=1}^N \beta_i x_{pi} + \epsilon_p \quad \text{Ecuación 9}$$

Donde $\epsilon_p \sim N(0, \sigma^2)$

El objetivo es encontrar los coeficientes β_i que minimizan la suma de cuadrados de los errores:

$$\Delta = \sum_{p=1}^N (y_p - \sum_{i=0}^N \beta_i x_{pi})^2 \quad \text{Ecuación 10}$$

Este problema es equivalente al recogido por una red neuronal con una sola capa, donde los parámetros β_i equivalen a los pesos de la misma y la función de activación es la identidad.

4.5 Perceptrón Simple

A partir del modelo de la neurona artificial Rosenblatt (1958,1962) desarrollo el modelo del Perceptrón el cuál básicamente consiste en una capa de neuronas con pesos y umbral ajustables como se muestra en la **Gráfico 14**, este sistema neuronal puede ser llamado una red neuronal debido a las conexiones existentes en sus elementos.

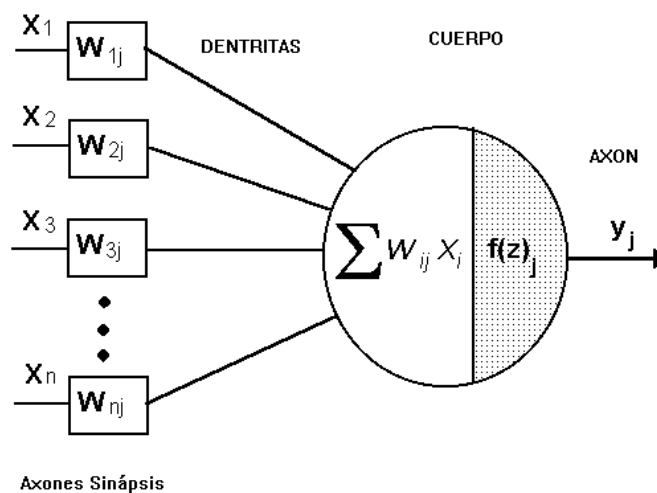


Gráfico 14: Esquema del Perceptrón simple.

Fue el mismo Rosenblatt quien en base al modelo de aprendizaje, determino el algoritmo de entrenamiento del Perceptrón que, siguiendo los principios de la regla delta, consiste en determinar el ajuste que se debe realizar a cada peso w en la neurona para que el error a la

salida sea cero. El algoritmo de entrenamiento del Perceptrón comienza por inicializar los parámetros libres (pesos y umbrales) en cero y posteriormente se estimula la red con algún vector de entrada obteniendo el error a cada salida como:

$$e_k(n) = d_k(n) - y_k(n) \quad \text{Ecuación 11}$$

Donde $e_k(n)$ corresponde al error a la salida para la muestra n de la neurona k , $d_k(n)$ es el valor destino esperado a la salida para la muestra n y $y_k(n)$ es la salida obtenida en la neurona k para la muestra n . El siguiente paso es calcular la delta o diferencia a agregar a los pesos para que el error $e_k(n)$ sea cero. La delta está dada por la siguiente ecuación:

$$\Delta w = e(n) \times (n) \quad \text{Ecuación 12}$$

$$\Delta \theta = e(n) \quad \text{Ecuación 13}$$

Donde $\Delta \theta$ es la modificación que se debe realizar al umbral y Δw es la modificación que se debe realizar a los pesos. La regla de aprendizaje del Perceptrón puede ser expresada como sigue:

$$w^{nuevo} = w^{actual} + e(n) \times (n) \quad \text{Ecuación 14}$$

$$\theta^{nuevo} = \theta^{actual} + e(n) \quad \text{Ecuación 15}$$

Si aplicamos esta regla a nuestro Perceptrón con distintas muestras, hasta que el error sea cero, se obtendrá una red que pueda generar exactamente las salidas deseadas para determinadas entradas logrando así que el Perceptrón aprenda una función.

Como se puede observar la regla delta para el Perceptrón consiste en agregar a cada peso el producto del error generado por la neurona con las entradas correspondientes de manera que se dan 3 casos posibles:

1. El error es cero y los pesos no se modifican.
2. El error es 1 y a cada peso se le agrega exactamente la entrada correspondiente.
3. El error es -1 y a cada peso se le subtrae exactamente la entrada correspondiente.

4.5.1 Limitaciones del perceptrón

Se debe tener en cuenta que no siempre el algoritmo de entrenamiento del Perceptrón podrá converger hacia un error nulo. De hecho, el Perceptrón es incapaz de converger en aquellas funciones que no son linealmente separables, es decir, aquellas cuyos elementos pueden ser separados por una línea recta.

Esto se debe a las propiedades inherentes de las unidades básicas del Perceptrón que son las neuronas artificiales, cuya limitación reside principalmente en la función de activación que como se puede observar separa las entradas en dos salidas de manera que con todas las entradas posibles se forma un hiperplano cuyos elementos de salida son separados en dos partes dentro del hiperplano.

Visto de otra forma, podríamos decir que el Perceptrón divide en dos grupos las entradas por medio de una línea divisoria de manera que no es posible separar elementos que no se encuentren claramente separados de otros elementos. Es decir que no se puede caracterizar elementos no lineales.

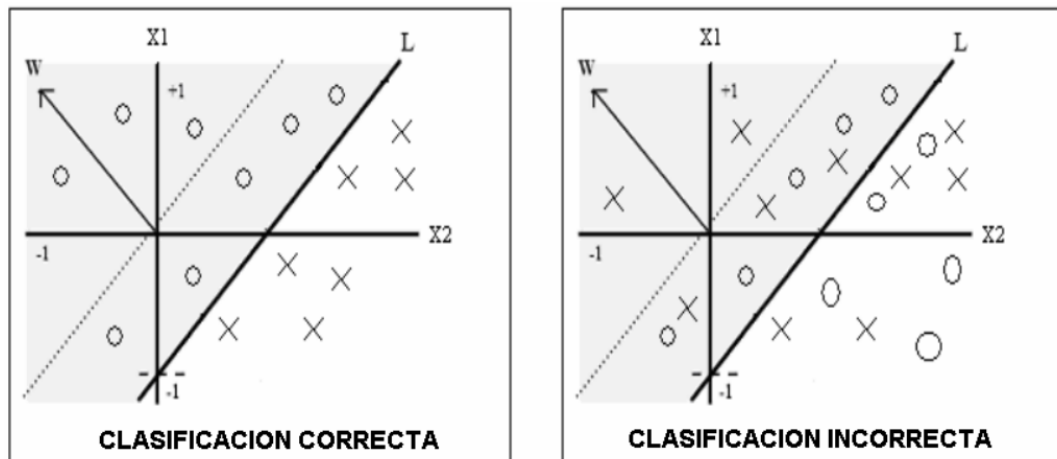


Gráfico 15: Margen de caracterización del perceptrón.

Esto se puede observar claramente en el **Gráfico 15** donde se muestra el caso en el que el Perceptrón es un buen clasificador, y otro caso donde el Perceptrón no puede

separar los elementos en dos categorías. En el gráfico, los puntos (X_1, X_2) corresponden a la intersección de dos entradas X_1 y X_2 .

Enfocándonos al problema de la detección de Intrusos, inferimos que el uso del Perceptrón sería inadecuado debido a las limitaciones lineales que su estructura posee. Sin embargo, a partir de estos principios se comenzará a vislumbrar las características necesarias para que nuestro sistema neuronal pueda clasificar patrones no lineales como los que se presentan comúnmente en una red de comunicaciones.

En el diseño de un sistema de detección de intrusos basado en sistemas neuronales es importante conocer los principios que llevan a este tipo de sistemas a realizar clasificaciones sobre vectores de entrada ya que nos ayudará a ubicar paulatinamente el tipo de diseño que pudiera ser el más adecuado para enfrentar el problema en cuestión.

Como se podrá ir viendo más adelante, el uso del paradigma de aprendizaje supervisado se presenta más adecuado para solucionar el problema de la clasificación binaria de datos y los distintos algoritmos que pudieran ajustar los parámetros libres de nuestra red, son todos basados en los principios asentados en este apartado por la regla delta. Únicamente cambiarán la función de costo a minimizar y la forma de minimizarla, sin embargo, en principio, el procedimiento será el mismo.

4.6 Perceptrón Multicapas.

El Perceptrón Multicapa es capaz de actuar como un aproximador universal de funciones. Esta propiedad convierte a las redes perceptrón multicapa en herramientas de propósito general, flexibles y no lineales.

Rumelhart et al. (1986) formalizaron un método para que una red del tipo perceptrón multicapa aprendiera la asociación que existe entre un conjunto de patrones de entrada y

sus salidas correspondientes: **método backpropagation-error** (propagación del error hacia atrás).

Una red con backpropagation, conteniendo al menos una capa oculta con suficientes unidades no lineales, puede aproximar cualquier tipo de función o relación continua entre un grupo de variables de entrada y salida.

Esta red tiene la capacidad de generalización: facilidad de dar estimaciones correctas de observaciones que no han sido incluidas en la fase de entrenamiento.

4.6.1 Arquitectura

Un perceptrón multicapa está compuesto por una capa de entrada, una capa de salida y una o más capas ocultas; aunque se ha demostrado que para la mayoría de problemas bastará con una sola capa oculta. En el **Gráfico 16** se puede observar un perceptrón típico formado por una capa de entrada, una capa oculta y una de salida.

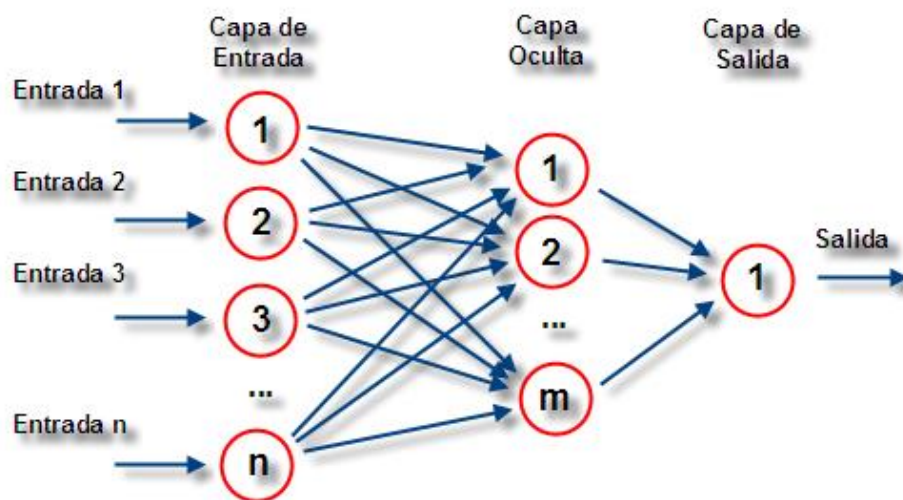


Gráfico 16: Esquema de un perceptrón multicapas.

Las conexiones entre neuronas son siempre hacia delante: las conexiones van desde las neuronas de una determinada capa hacia las neuronas de la siguiente capa; no

hay conexiones laterales ni conexiones hacia atrás. Por tanto, la información siempre se transmite desde la capa de entrada hacia la capa de salida.

Como notación se denomina w_{ji} al peso de conexión entre la neurona de entrada i y la neurona oculta j , y v_{kj} al peso de conexión entre la neurona oculta j y la neurona de salida k .

4.6.2 Algoritmo backpropagation

Se considera una etapa de funcionamiento donde se presenta un patrón de entrada y este se transmite a través de las sucesivas capas de neuronas hasta obtener una salida.

Después, hay una etapa de entrenamiento o aprendizaje donde se modifican los pesos de la red de manera que coincida la salida objetivo con la salida obtenida por la red.

4.6.3 Etapa de funcionamiento

Cuando se presenta un patrón p de entrada $X^p: x^p_1, \dots, x^p_i, \dots, x^p_N$, este se transmite a través de los pesos w_{ji} desde la capa de entrada hacia la capa oculta. Las neuronas de esta capa intermedia transforman las señales recibidas mediante la aplicación de una función de activación proporcionando, de este modo, un valor de salida. Este se transmite a través de los pesos v_{kj} hacia la capa de salida, donde aplicando la misma operación que en el caso anterior, las neuronas de esta última capa proporcionan la salida de la red.

Este proceso se resume en lo siguiente:

La entrada total, θ^p_j , que recibe una neurona oculta j es:

$$\theta^p_j = \sum_{i=1}^N w_{ji}x^p_i + \lambda_i \quad \text{Ecuación 16}$$

Donde λ_i es un peso asociado a una neurona ficticia con valor de salida igual a 1 que hace el papel de término independiente o *intercept*.

El valor de salida de la neurona oculta j , y_j^p , se obtiene aplicando una función de activación $f(\cdot)$ sobre su entrada neta:

$$y_j^p = f(\theta_j^p) \quad \text{Ecuación 17}$$

De igual forma, la entrada neta que recibe una neurona de salida k , θ_k^p es:

$$\theta_k^p = \sum_{j=1}^H v_{kj} x_j^p + \lambda_k \quad \text{Ecuación 18}$$

Por último, el valor de salida de la neurona de salida k , y_k^p es:

$$y_k^p = f(\theta_k^p) \quad \text{Ecuación 19}$$

4.6.4 Etapa de aprendizaje

En la etapa de aprendizaje, el objetivo es hacer mínimo el error entre la salida obtenida por la red y la salida óptima ante la presentación de un conjunto de patrones, denominado grupo de entrenamiento. Así, el aprendizaje en las redes backpropagation es de tipo supervisado.

La función de error que se pretende minimizar para cada patrón p viene dada por:

$$E^p = \frac{1}{2} \sum_{k=1}^N (d_k^p - y_k^p)^2 \quad \text{Ecuación 20}$$

Donde d_k^p es la salida deseada para la neurona de salida k ante la presentación del patrón p .

A partir de esta expresión se puede obtener una medida general del error total mediante:

$$E = \sum_{p=1}^P E^p \quad \text{Ecuación 21}$$

La base del algoritmo backpropagation para la modificación de los pesos es la técnica conocida como **gradiente decreciente**.

Como E^p es función de todos los pesos de la red, el gradiente de E^p es un vector igual a la derivada parcial de E^p respecto de cada uno de los pesos.

El gradiente toma la dirección que determina el incremento más rápido en el error, mientras que la dirección opuesta, es decir, la dirección negativa, determina el decremento más rápido en el error.

Por tanto, el error puede reducirse ajustando cada peso en esa dirección:

$$-\sum_{p=1}^P \frac{\partial E^p}{\partial w_{ji}} \quad \text{Ecuación 22}$$

Un peligro que puede surgir al utilizar el método del gradiente decreciente es que el aprendizaje converja a un mínimo local. Sin embargo, el problema potencial de los mínimos locales se da en raras ocasiones en datos reales.

A nivel práctico, la forma de modificar los pesos de forma iterativa consiste en aplicar la regla de la cadena a la expresión del gradiente y añadir una tasa dada de aprendizaje η . Así, en una neurona de salida se tendría:

$$\Delta v_{kj}(n+1) = -\eta \frac{\partial E^p}{\partial v_{ki}} = \eta \sum_{p=1}^P \delta_k^p y_j^p \quad \text{Ecuación 23}$$

Donde

$$\delta_k^p = (d_k^p - y_k^p) f'(\theta_k^p) \quad \text{Ecuación 24}$$

Y n indica la iteración.

En una neurona oculta:

$$\Delta w_{ji}(n+1) = \eta \sum_{p=1}^P \delta_j^p y_i^p \quad \text{Ecuación 25}$$

Donde:

$$\delta_k^p = f(\theta_j^p) \sum_{k=1}^M \delta_k^p v_{kj}^p \quad \text{Ecuación 26}$$

Se puede observar que el error o valor asociado a una neurona oculta j , viene determinado por la suma de los errores que se cometen en las k neuronas de salida que reciben como entrada la salida de esa neurona oculta j . De ahí que el algoritmo también se denomine *propagación del error hacia atrás*.

Para la modificación de los pesos, la actualización se realiza después de haber presentado todos los patrones de entrenamiento. Este es el modo habitual de proceder y se denomina aprendizaje por *lotes* o modo *batch*.

Otra modalidad denominada aprendizaje *en serie* o modo *on-line* consiste en actualizar los pesos tras la presentación de cada patrón de entrenamiento que ha de hacerse en orden aleatorio.

Para acelerar el proceso de convergencia de los pesos, Rumelhart et al. (1986) sugirieron añadir un término, denominado **momento**, que tiene en cuenta la dirección del incremento tomada en la iteración anterior:

$$\Delta v_{kj}(n+1) = \eta \left(\sum_{p=1}^P \delta_k^p y_j^p \right) + \alpha \Delta v_{kj}(n) \quad \text{Ecuación 27}$$

4.6.5 Fases en la aplicación de un perceptrón multicapa

Una red del tipo perceptrón multicapa intenta resolver dos tipos de problemas:

- **Problemas de predicción**, que consisten en la estimación de una variable continua de salida, a partir de la presentación de un conjunto de variables predictoras de entrada (discretas y/o continuas).
- **Problemas de clasificación**, que consisten en la asignación de la categoría de pertenencia de un determinado patrón a partir de un conjunto de variables predictoras de entrada (discretas y/o continuas).

4.6.6 Selección de las variables relevantes y preprocesamiento de los datos

Para obtener una buena aproximación, se deben elegir cuidadosamente las variables a emplear: se trata de incluir en el modelo las variables predictoras que realmente predigan la variable dependiente o de salida, pero que a su vez no tengan relaciones entre sí, ya que esto puede provocar un sobreajuste innecesario en el modelo.

Las variables deben seguir una distribución normal o uniforme, y el rango de posibles valores debe ser aproximadamente el mismo y acotado dentro del intervalo de trabajo de la función de activación empleada en las capas ocultas y de salida de la red neuronal.

Así, las variables de entrada y salida suelen acotarse en valores comprendidos entre 0 y 1 o entre -1 y 1.

Si la variable es discreta, se utiliza la codificación dummy.

Por ejemplo, la variable sexo podría codificarse como: 0 = hombre, 1 = mujer; estando representada por una única neurona.

La variable nivel social podría codificarse como: 100 = bajo, 010 = medio, 001 = alto; estando representada por tres neuronas.

Por su parte, si la *variable es de naturaleza continua*, ésta se representa mediante una sola neurona, como, por ejemplo, la pensión de una persona.

4.6.7 Entrenamiento de una neurona

a. Elección de los pesos iniciales

Se hace una asignación de pesos pequeños generados de forma aleatoria en un rango de valores entre -0.5 y 0.5 o algo similar.

b. Arquitectura de la red

Respecto a la arquitectura de la red, se sabe que para la mayoría de problemas prácticos bastará con utilizar una sola capa oculta.

El número de neuronas de la capa de entrada está determinado por el número de variables predictoras.

Así, en los ejemplos anteriores, la variable sexo estaría representada por una neurona que recibiría los valores 0 o 1. La variable estatus social estaría

representada por tres neuronas. La variable pensión de una persona estaría representada por una neurona que recibiría un valor previamente acotado, por ejemplo, a valores entre 0 y 1.

El número de neuronas de la capa de salida está determinado según el mismo esquema que en el caso anterior.

Cuando intentamos discriminar entre dos categorías, bastará con utilizar una única neurona. Por ejemplo, salida 1 para la categoría A, salida 0 para la categoría B.

Si estamos ante un problema de estimación de una variable continua, tendremos una única neurona que dará como salida el valor de la variable a estimar.

El número de neuronas ocultas determina la capacidad de aprendizaje de la red neuronal. Para evitar el sobreajuste, se debe usar el mínimo número de neuronas ocultas con las cuales la red funcione de forma adecuada. Esto se consigue evaluando el rendimiento de diferentes arquitecturas en función de los resultados obtenidos con el grupo de validación.

4.6.8 Tasa de aprendizaje y factor momento

El valor de la tasa de aprendizaje (η) controla el tamaño del cambio de los pesos en cada iteración. Se deben evitar dos extremos:

- un *ritmo de aprendizaje demasiado pequeño* puede ocasionar una disminución importante en la velocidad de convergencia y la posibilidad de acabar atrapado en un mínimo local;

- en cambio, un *ritmo de aprendizaje demasiado grande* puede conducir a inestabilidades en la función de error, lo cual evitará que se produzca la convergencia debido a que se darán saltos en torno al mínimo sin alcanzarlo.

Por tanto, se recomienda elegir un ritmo de aprendizaje lo más grande posible sin que provoque grandes oscilaciones. En general, el valor de la tasa de aprendizaje suele estar comprendida entre 0.05 y 0.5.

El factor momento (α) acelera la convergencia de los pesos. Se suele tomar un valor próximo a 1 (por ejemplo, 0.9).

4.6.9 Función de activación de las neuronas ocultas y de salida

Se usan dos funciones básicas: la función lineal (o identidad), la función signo y funciones sigmoideas (como la función logística o la función tangente hiperbólica).

En general, se utiliza una función sigmoidea como función de activación en las neuronas de la capa oculta.

La elección de la función de activación en las neuronas de la capa de salida dependerá del tipo de tarea que se considera:

- En *tareas de clasificación*, se toma la función de activación sigmoidea.
- En cambio, en *tareas de predicción o aproximación de una función*, generalmente se toma la función de activación lineal.

4.6.10 Evaluación del rendimiento del modelo

Una vez seleccionado el modelo de red que ha obtenido el mejor resultado con el conjunto de validación, se debe evaluar la capacidad de generalización de la red con otro grupo de datos independientes, o conjunto de datos de test.

Se utiliza la media cuadrática del error para evaluar el modelo:

$$MC_{error} = \frac{\sum_{p=1}^P \sum_{k=1}^M (a_k^p - y_k^p)^2}{P \times M}$$

En problemas de clasificación de patrones es mejor usar el *porcentaje de clasificaciones correctas e incorrectas*. Se puede construir una **tabla (matriz) de confusión** y calcular diferentes índices de *asociación* y acuerdo entre el criterio y la decisión tomada por la red neuronal.

4.6.11 Interpretación de los pesos obtenidos

Se trata de interpretar los pesos de la red neuronal. El método más popular es el análisis de sensibilidad.

El análisis de sensibilidad está basado en la medición del efecto que se observa en una salida y_k debido al cambio que se produce en una entrada x_i . Cuanto mayor efecto se observe sobre la salida, mayor sensibilidad se puede deducir que presenta respecto a la entrada.

Un método común consiste en fijar el valor de todas las variables de entrada a su valor medio e ir variando el valor de una de ellas a lo largo de todo su rango, registrando el valor de salida de la red. (Russell & Norvig, *Inteligencia Artificial, un enfoque moderno*, 2004).

PARTE
EXPERIMENTAL

1. EQUIPOS Y REACTIVOS

1.1 Reactivos

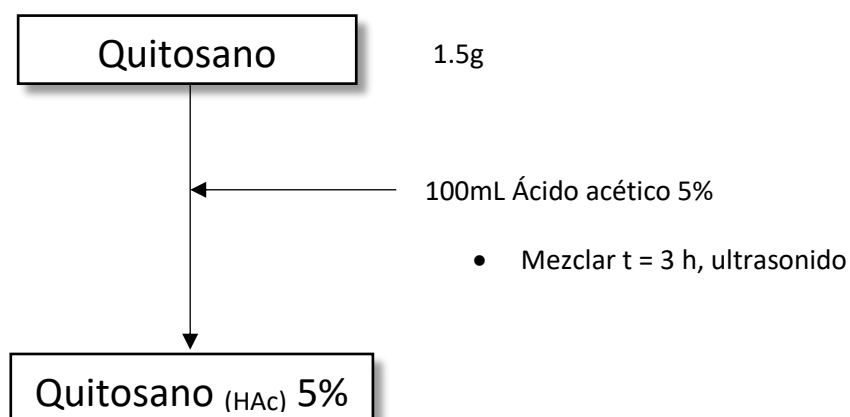
- Estándar de cobre marca MERCK de 1000 ppm
- Quitosano en polvo de alto peso molecular, marca Sigma Aldrich
- Glutaraldehído al 50%, marca Sigma Aldrich
- Ácido acético 99.8%, marca Fermont
- Hidróxido de sodio en presentación granallas, marca Merck.
- Ácido clorhídrico concentrado, marca Fermont

1.2 Equipos

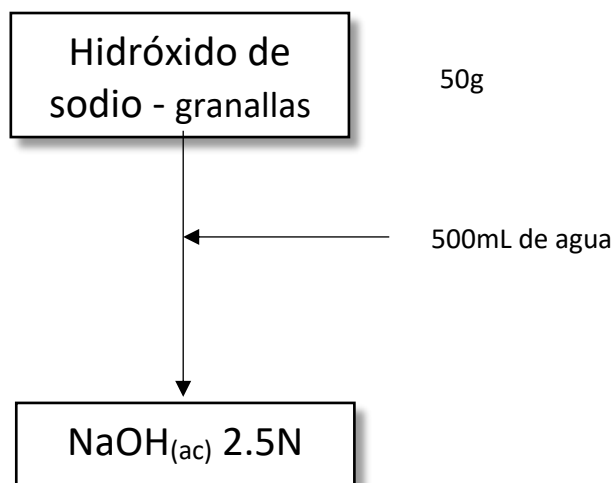
- Espectrofotómetro de Absorción Atómica Varian AA240
- Medidor de pH Thermo Scientific Orion, Orion Star

2. PROCEDIMIENTO EXPERIMENTAL

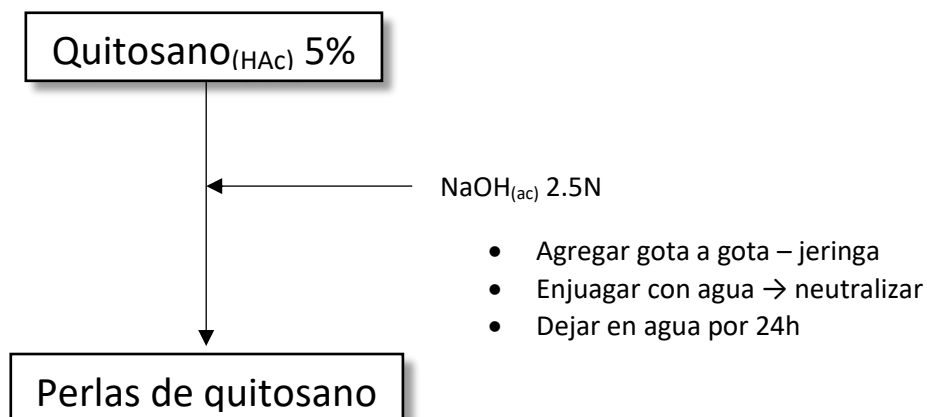
2.1 Preparación de quitosano 5%



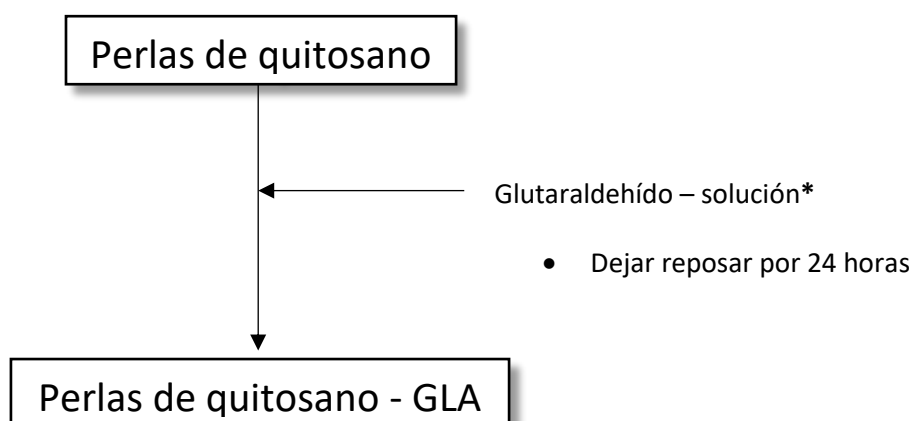
2.2 Preparación de hidróxido de sodio 2.5N



2.3 Preparación de perlas de quitosano

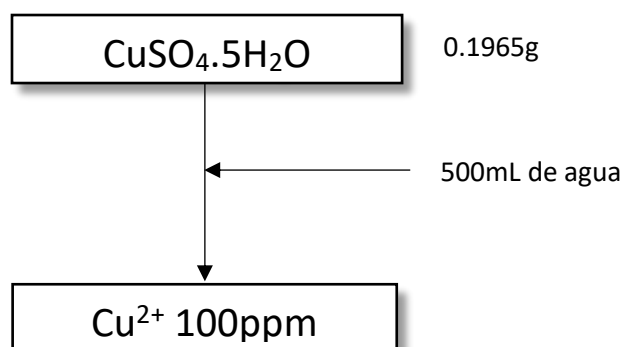


2.4 Entrecruzamiento de quitosano con glutaraldehído



*: La cantidad depende de la matriz de diseños de experimentos (**Tabla 7**).

2.5 Preparación de solución de cobre 100 ppm



Adicional:

Se requerirá soluciones de Cu^{2+} (dependiendo del diseño de experimentos), HCl 0.01N y NaOH 0.01N.

2.6 Procedimiento experimental

Se considera como la tabla de niveles de experimentos:

Tabla 7: Tabla de niveles experimentales

Factor	Nivel inferior	Nivel superior
pH	3	5
Masa de bp (g)	0.5	1
Conc GLA (g/100mL)	1	2
Tiempo (min)	3	5
Agitación (rpm)	300	500
Conc inicial Cu (mg/L)	10	20

Se realizó, en base a la matriz de factores y niveles (**Tabla 7**), para la etapa de entrenamiento, un diseño factorial completo de 6 niveles, considerando 5 puntos centrales, para todos los factores estudiados, a excepción de la concentración de GLA, definiendo así un DOE: 2^6 (+5PC). El sistema se compone por 50 mL de solución de Cu.

En base a ello, se tiene en la **Tabla 8** la matriz de diseño de experimentos:

Tabla 8: Matriz de diseño de experimentos

N°	pH	masa bp (g)	conc GLA (g/100mL)	tiempo (min)	agitación (rpm)	conc Cu inicial(mg/L)
1	3	0.5	1	3	100	10
2	5	0.5	1	3	100	10
3	3	1	1	3	100	10
4	5	1	1	3	100	10
5	3	0.5	2	3	100	10
6	5	0.5	2	3	100	10
7	3	1	2	3	100	10
8	5	1	2	3	100	10
9	3	0.5	1	5	100	10
10	5	0.5	1	5	100	10
11	3	1	1	5	100	10
12	5	1	1	5	100	10
13	3	0.5	2	5	100	10
14	5	0.5	2	5	100	10
15	3	1	2	5	100	10
16	5	1	2	5	100	10
17	3	0.5	1	3	300	10
18	5	0.5	1	3	300	10
19	3	1	1	3	300	10
20	5	1	1	3	300	10
21	3	0.5	2	3	300	10
22	5	0.5	2	3	300	10
23	3	1	2	3	300	10
24	5	1	2	3	300	10
25	3	0.5	1	5	300	10
26	5	0.5	1	5	300	10
27	3	1	1	5	300	10
28	5	1	1	5	300	10
29	3	0.5	2	5	300	10
30	5	0.5	2	5	300	10
31	3	1	2	5	300	10
32	5	1	2	5	300	10
33	3	0.5	1	3	100	20
34	5	0.5	1	3	100	20
35	3	1	1	3	100	20

36	5	1	1	3	100	20
37	3	0.5	2	3	100	20
38	5	0.5	2	3	100	20
39	3	1	2	3	100	20
40	5	1	2	3	100	20
41	3	0.5	1	5	100	20
42	5	0.5	1	5	100	20
43	3	1	1	5	100	20
44	5	1	1	5	100	20
45	3	0.5	2	5	100	20
46	5	0.5	2	5	100	20
47	3	1	2	5	100	20
48	5	1	2	5	100	20
49	3	0.5	1	3	300	20
50	5	0.5	1	3	300	20
51	3	1	1	3	300	20
52	5	1	1	3	300	20
53	3	0.5	2	3	300	20
54	5	0.5	2	3	300	20
55	3	1	2	3	300	20
56	5	1	2	3	300	20
57	3	0.5	1	5	300	20
58	5	0.5	1	5	300	20
59	3	1	1	5	300	20
60	5	1	1	5	300	20
61	3	0.5	2	5	300	20
62	5	0.5	2	5	300	20
63	3	1	2	5	300	20
64	5	1	2	5	300	20
65	4	0.75	1	4	200	15
66	4	0.75	2	4	200	15
67	4	0.75	1	4	200	15
68	4	0.75	2	4	200	15
69	4	0.75	1	4	200	15
70	4	0.75	2	4	200	15
71	4	0.75	1	4	200	15
72	4	0.75	2	4	200	15
73	4	0.75	1	4	200	15
74	4	0.75	2	4	200	15

2.7 Definición de la red neuronal artificial

En la presente investigación, para el desarrollo de la red neuronal artificial propuesta se hace uso de la herramienta *Network Neural Toolbox* (NNTOOL), paquete aplicativo del software *MATLAB R2014a*. Para esta investigación se realizó un perceptrón multicapa del tipo *feed forward neural network* (FFNN) con *Back-Propagation* (BP). Se probó el número de capas ocultas necesarias para ajustar el modelo a los datos obtenidos (esto es, el número de capas ocultas que minimice el MSE), desde 1 a 20 capas ocultas. Para la función de activación entre la capa de entrada y la capa oculta se utilizó la tangente hiperbólica (*tansing*) y como función de transferencia entre la capa oculta y la capa de salida, la tangente hiperbólica (*tansing*). Las relaciones son como sigue:

$$tansing(suma) = \frac{1 - \exp(-suma)}{1 + \exp(-suma)} \quad \text{Ecuación 29}$$

Para evitar una conglomeración en los resultados, las variables de entrada y salida se normalizaron dentro de un intervalo uniforme [0; 1] de acuerdo con la siguiente ecuación:

$$x_{norm} = \frac{(x - x_{min})}{x_{max} - x_{min}} \quad \text{Ecuación 30}$$

Donde x es la variable a normalizar, x_{max} es el valor máximo valor y x_{min} es el mínimo valor, de la serie de resultados de los patrones de aprendizaje.

RESULTADOS OBTENIDOS Y DISCUSIONES

Los resultados obtenidos al realizar los ensayos de la matriz de experimentos (**Tabla 8**) se encuentran tabulados en la **Tabla 11** en la columna “% **Remoción**”.

Estos datos (entrenamiento) ingresan como patrones en la red neuronal, y se va haciendo modificación en el número de capas ocultas entre 1 a 20, buscando cual es el número mínimo en el cual se tiene un Error cuadrático medio (Mean Square Error: MSE) que se puede aceptar, teniendo fijado los siguientes parámetros:

Tabla 9: Valores fijados en la elaboración de la RNA.

<i>Parámetro</i>	<i>valor</i>
<i>network Type</i>	Feed-forward backpropagation
<i>training function</i>	TRAINGD
<i>NUMBER OF LAYERS</i>	02
<i>TRANSFER FUNCTION (LAYER 1)</i>	TANSIG
<i>TRANSFER FUNCTION (LAYER 2)</i>	TANSIG

Para ello el modelo propone los siguientes errores dependiendo de la cantidad de capas ocultas:

$$MSE = \frac{\sum_{i=1}^N (Y_t - Y_N)^2}{N} \quad \text{Ecuación 31}$$

Tabla 10: MSE determinado por cada RNA, con diferentes números de capas ocultas (por WEKA v.3.6).

N° Capas ocultas	MSE
0 (perceptrón simple)	64.93%
1	67.40%
2	55.38%
3	38.63%
4	21.16%
5	20.31%
6	19.88%
7	17.03%
8	14.62%
9	16.44%
10	17.61%
11	18.79%

12	17.51%
13	18.35%
14	14.12%
15	11.23%
16	10.64%
17	10.52%
18	10.39%
19	9.93%
20	6.80%

Como se evidencia, a partir de 16 capas ocultas se puede obtener un %MSE aceptable (10% de error en el entrenamiento), por lo que fijamos este número como el número de capas ocultas a trabajar (**Gráfico 17**). Se pudo elegir la RNA con 20 capas ocultas, incluso se esperaría que al aumentar a más números de capas ocultas el error de predicción sea mejor, sin embargo, se busca el número mínimo de capas para obtener un error aceptable por el tema de optimizar recursos computacionales, pues a mayor número de capas, la red demora más en converger el error del modelo (que se busca siempre aproximarse lo mas posible a 0%).

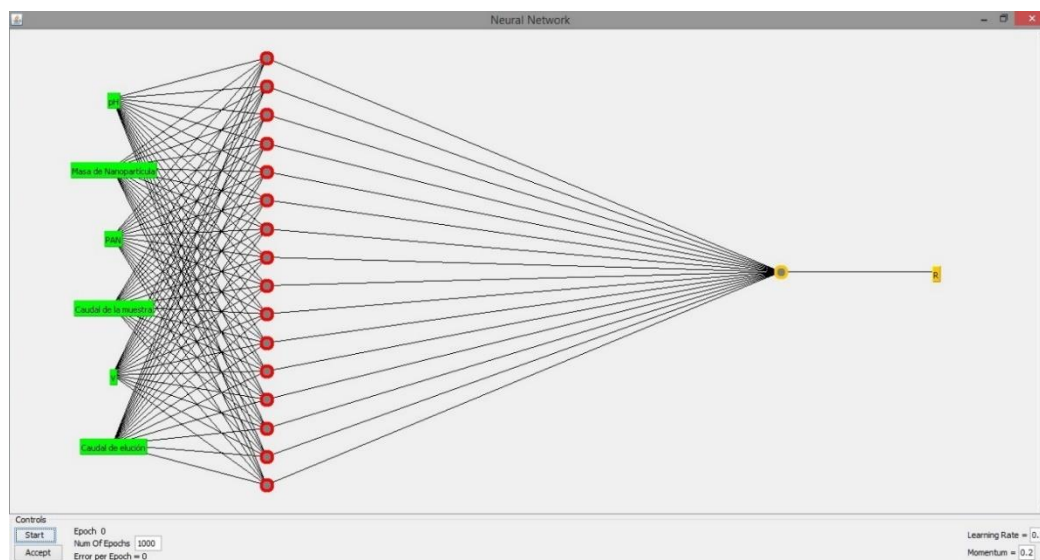


Gráfico 17: RNA con 16 capas ocultas (6 – 16 – 1). Gráfico proporcionado en WEKA.

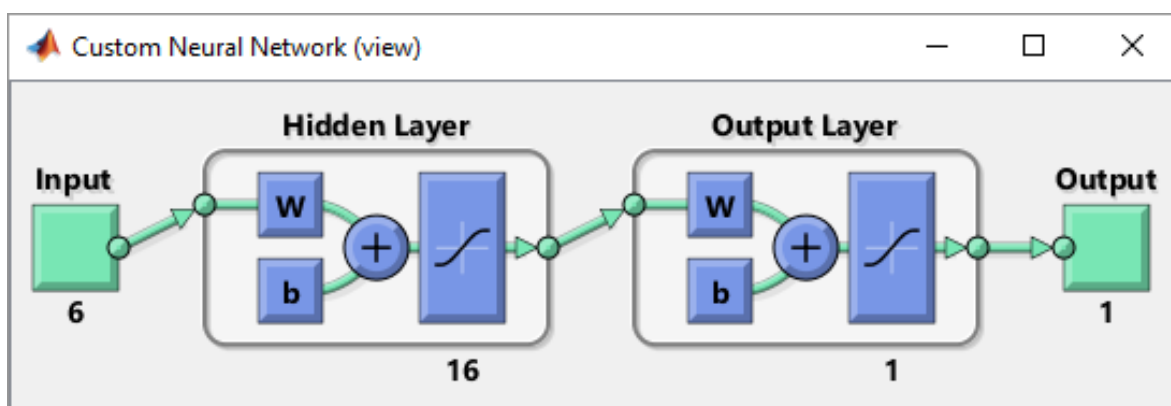


Gráfico 18: RNA con 16 capas ocultas (6 – 16 – 1). Gráfico proporcionado en MATLAB.

Los parametros de ingresos para esta red neuronal artificial se muestran:

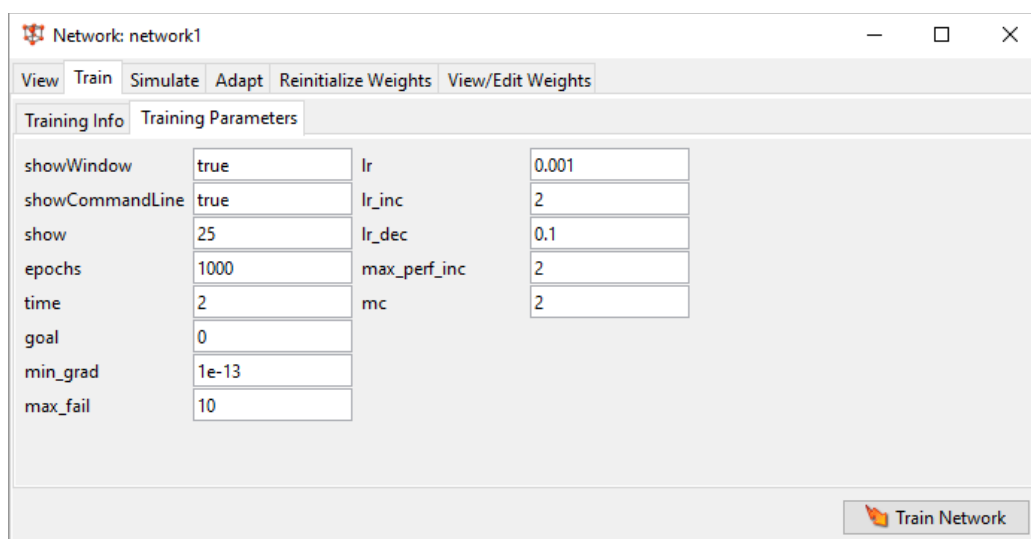


Gráfico 19: Parámetros para entrenamiento de la RNA (6 – 16 – 1).

Se procede a realizar la etapa de entrenamiento (*training*) a la red neuronal artificial (6 – 16 – 1, **Gráfico 18**) con los parámetros anteriores (**Gráfico 19**), teniendo la convergencia en los valores siguientes (**Gráfico 20**):

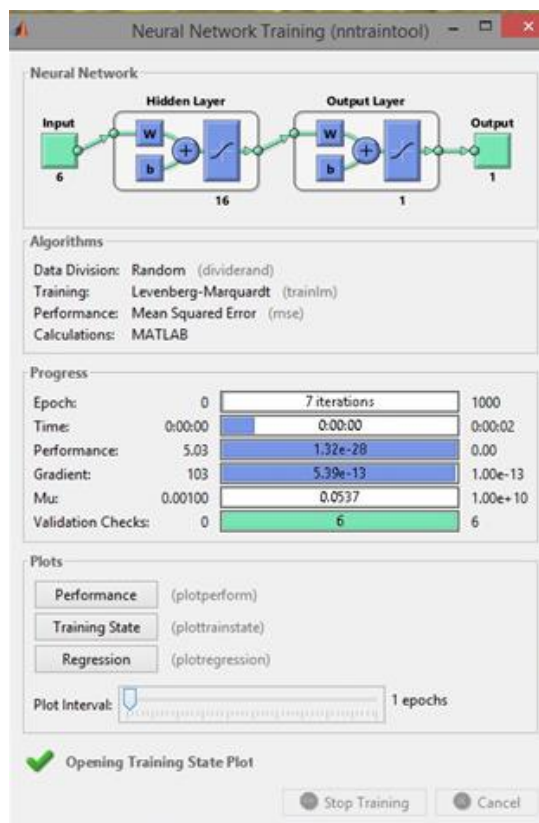


Gráfico 20: Fases y resultados del entrenamiento de la RNA (6 – 16 – 1).

Se tiene el resumen gráfico del proceso simultáneo de entrenamiento, validación y prueba (*training - validation - test*, los cuales están estipulados por MATLAB como 70% - 15% - 15% de la data general: 74 datos).

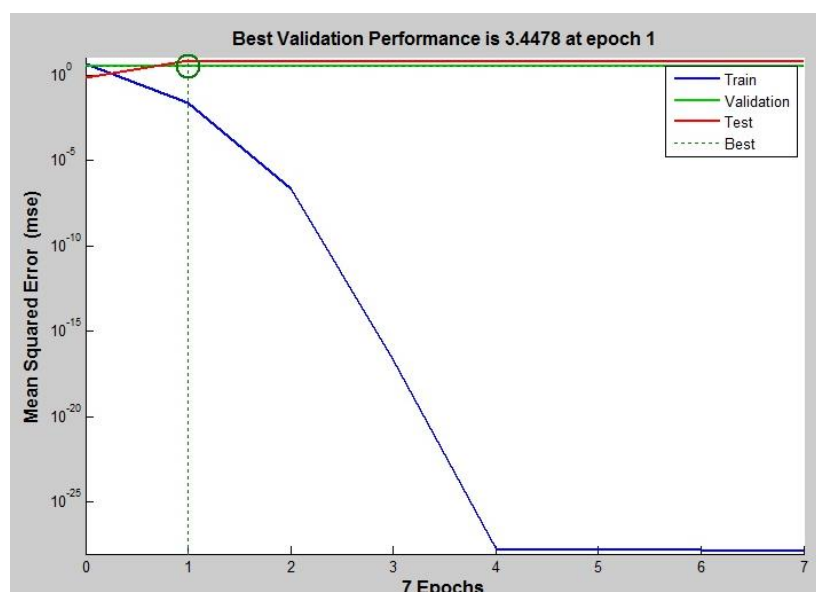


Gráfico 21: Perfil del rendimiento de la validación de la RNA (6 – 16 – 1).

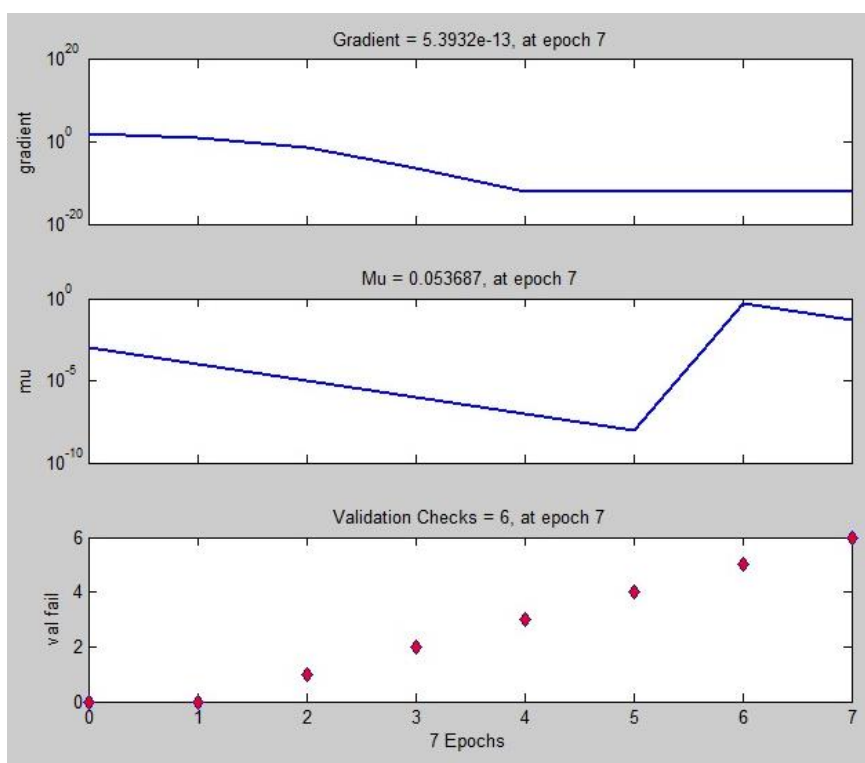


Gráfico 22: Perfil de evolución de los parámetros de la RNA por épocas de aprendizaje.

Por el **Gráfico 21** se ve que el mejor rendimiento en la validación del modelo de la RNA (6 – 16 – 1) es en la primera época, donde alcanza el valor de 3.44% de error. En el **Gráfico 22** se puede ver que la gradiente de la función de correspondencia experimental – predictiva, toma un valor mínimo constante en la época 4, llegando a la época 7 un valor de gradiente igual a 5.3932×10^{-13} (el valor de detención de iteración del **Gráfico 19** es de 1×10^{-13}). El momento máximo de obtiene en la época 6, sin embargo todos los procesos se detienen en la época de aprendizaje $n = 6$ pues ya hay 6 valores que han salido fuera de control en la validación / prueba de la RNA (**Gráfico 20**), pero pese a llegar al límites de predicciones fuera de control, la gradiente ha llegado a un valor bastante cercano al valor esperado, razón por la cual se acepta el modelo de predicción presentado: RNA (6 – 16 – 1).

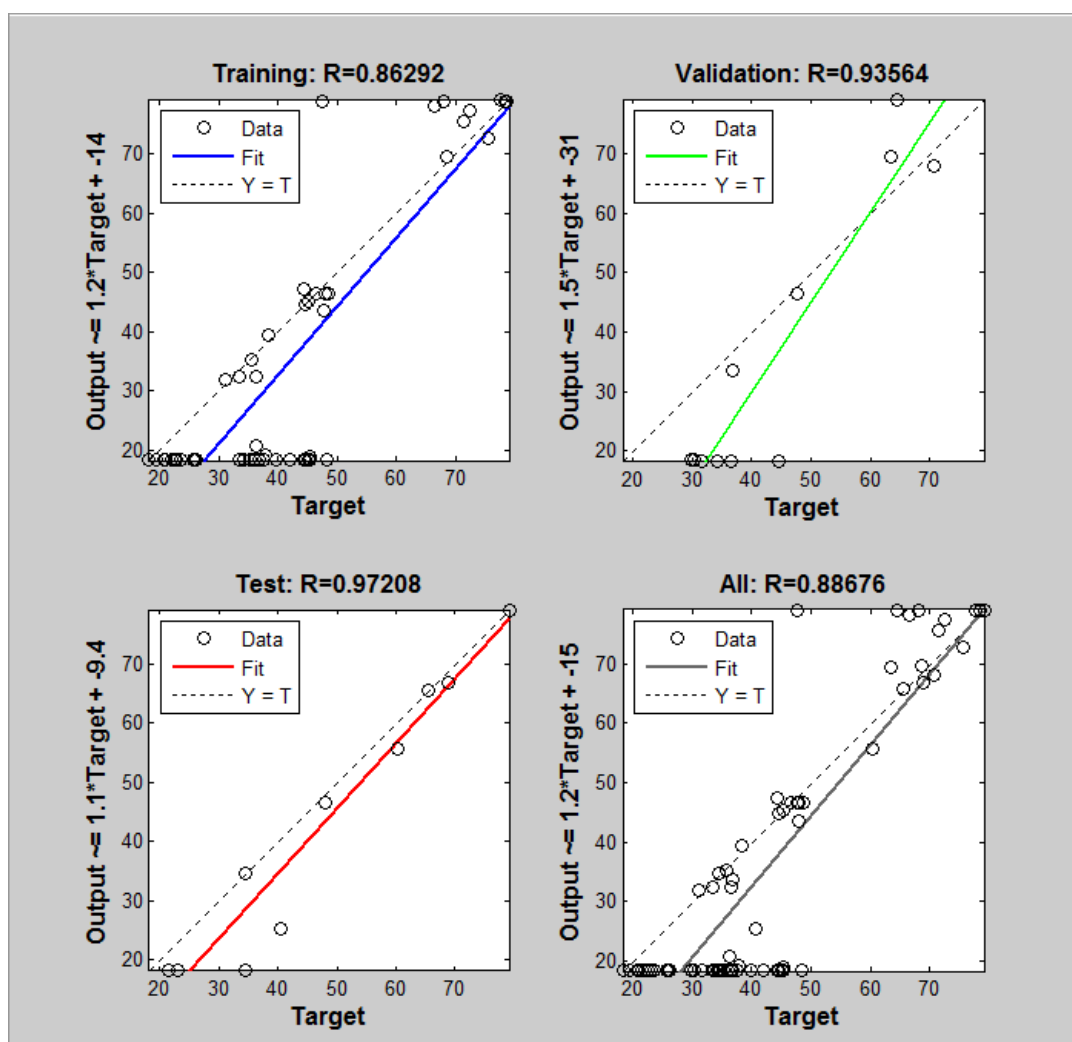


Gráfico 23: Gráfica de correlación en la fase de entrenamiento, validación, prueba.

La correlación de respuesta dada de la RNA (en MATLAB) es de 0.88676 (**Gráfico 23**), por lo que se ve que hay una correlación lineal aceptable. El modelo es predictorio a un $0.88676^2 = 0.7863 = 78.63\%$ con el modelo de linealidad (para los 74 patrones), lo cual indica que la RNA ha logrado un muy buen objetivo predictorio. Se ve además que para la etapa de entrenamiento, los 52 datos (70% de los datos) hay una correspondencia del $r = 0.86292$, por lo que la RNA ha aprendido con mucha eficiencia. Para la etapa de validación, los 11 datos elegidos al azar (15% de los datos) han tenido buen resultado predictorio ($r = 0.93564$), igualmente para la etapa de prueba (15% de datos al azar), el cual ha logrado un $r = 0.97208$. Cabe resaltar que es el MALTAB quien ha dividido el

conjunto de 74 patrones en: 52 datos para el entrenamiento (70%), 11 para la validación (15%) y 11 para la prueba (15%). Esta elección de MATLAB a sido aleatoria, evitando sesgo en el proceso.

Los datos obtenidos de predicción para los datos de entrenamiento, prueba y validación son como siguen (**Tabla 11**):

Tabla 11: Resultados obtenidos de la RNA entrenada, con 16 capas ocultas y con los parámetros fijados en la **Tabla 9** y **Gráfico 18**.

N°	pH	masa bp (g)	conc GLA (g/100mL)	tiempo (min)	agitación (rpm)	conc Cu inicial (mg/L)	Conc Final	% Remoción	%R_pred	%E
1	3	0.5	1	3	100	10	6.38	36.21	37.18	2.68%
2	5	0.5	1	3	100	10	6.17	38.34	40.39	5.36%
3	3	1	1	3	100	10	3.97	60.26	61.86	2.66%
4	5	1	1	3	100	10	3.54	64.58	65.46	1.36%
5	3	0.5	2	3	100	10	5.47	45.26	45.96	1.55%
6	5	0.5	2	3	100	10	5.24	47.59	47.93	0.73%
7	3	1	2	3	100	10	2.94	70.57	69.85	1.02%
8	5	1	2	3	100	10	2.14	78.57	74.97	4.58%
9	3	0.5	1	5	100	10	6.65	33.54	34.55	2.99%
10	5	0.5	1	5	100	10	6.21	37.85	38.18	0.87%
11	3	1	1	5	100	10	3.45	65.55	68.50	4.51%
12	5	1	1	5	100	10	3.15	68.50	69.18	0.99%
13	3	0.5	2	5	100	10	5.54	44.59	48.04	7.74%
14	5	0.5	2	5	100	10	5.21	47.87	49.48	3.37%
15	3	1	2	5	100	10	2.45	75.54	73.53	2.66%
16	5	1	2	5	100	10	2.09	79.10	78.82	0.35%
17	3	0.5	1	3	300	10	6.25	37.55	36.96	1.55%
18	5	0.5	1	3	300	10	5.80	41.95	42.71	1.80%
19	3	1	1	3	300	10	3.30	67.02	80.99	20.86%
20	5	1	1	3	300	10	3.20	67.99	57.72	15.10%
21	3	0.5	2	3	300	10	5.94	40.59	44.17	8.82%
22	5	0.5	2	3	300	10	5.56	44.41	50.49	13.69%
23	3	1	2	3	300	10	2.85	71.54	71.60	0.08%
24	5	1	2	3	300	10	2.24	77.59	76.18	1.82%
25	3	0.5	1	5	300	10	6.48	35.20	42.62	21.09%
26	5	0.5	1	5	300	10	6.02	39.81	39.25	1.39%
27	3	1	1	5	300	10	3.65	63.54	63.85	0.49%
28	5	1	1	5	300	10	3.12	68.76	70.32	2.27%
29	3	0.5	2	5	300	10	5.47	45.35	48.85	7.73%

30	5	0.5	2	5	300	10	5.16	48.41	48.76	0.73%
31	3	1	2	5	300	10	2.75	72.54	71.04	2.06%
32	5	1	2	5	300	10	2.15	78.54	79.70	1.47%
33	3	0.5	1	3	100	20	15.79	21.05	24.32	15.54%
34	5	0.5	1	3	100	20	15.37	23.16	25.08	8.31%
35	3	1	1	3	100	20	13.78	31.09	30.85	0.74%
36	5	1	1	3	100	20	12.87	35.67	32.43	9.09%
37	3	0.5	2	3	100	20	16.35	18.26	15.68	14.12%
38	5	0.5	2	3	100	20	15.51	22.44	19.50	13.10%
39	3	1	2	3	100	20	13.09	34.55	36.46	5.54%
40	5	1	2	3	100	20	12.71	36.44	38.76	6.39%
41	3	0.5	1	5	100	20	15.84	20.81	22.31	7.22%
42	5	0.5	1	5	100	20	15.26	23.69	24.81	4.74%
43	3	1	1	5	100	20	13.96	30.18	32.33	7.13%
44	5	1	1	5	100	20	13.13	34.34	35.20	2.51%
45	3	0.5	2	5	100	20	15.41	22.93	21.84	4.74%
46	5	0.5	2	5	100	20	14.82	25.92	24.88	4.03%
47	3	1	2	5	100	20	13.32	33.41	35.94	7.56%
48	5	1	2	5	100	20	12.64	36.79	37.32	1.42%
49	3	0.5	1	3	300	20	16.09	19.53	17.34	11.23%
50	5	0.5	1	3	300	20	15.38	23.12	20.54	11.13%
51	3	1	1	3	300	20	14.03	29.84	33.76	13.13%
52	5	1	1	3	300	20	13.08	34.60	35.71	3.20%
53	3	0.5	2	3	300	20	15.58	22.11	24.26	9.71%
54	5	0.5	2	3	300	20	14.76	26.20	24.57	6.24%
55	3	1	2	3	300	20	13.33	33.36	34.87	4.54%
56	5	1	2	3	300	20	12.83	35.87	42.72	19.09%
57	3	0.5	1	5	300	20	15.70	21.48	21.86	1.76%
58	5	0.5	1	5	300	20	15.59	22.04	24.06	9.15%
59	3	1	1	5	300	20	13.14	34.29	28.39	17.21%
60	5	1	1	5	300	20	12.79	36.06	38.54	6.87%
61	3	0.5	2	5	300	20	15.27	23.64	22.47	4.98%
62	5	0.5	2	5	300	20	14.87	25.65	30.23	17.87%
63	3	1	2	5	300	20	13.68	31.60	38.54	21.94%
64	5	1	2	5	300	20	12.63	36.83	42.31	14.89%
65	4	0.75	1	4	200	15	8.18	45.43	45.66	0.48%
66	4	0.75	2	4	200	15	7.71	48.59	47.05	3.16%
67	4	0.75	1	4	200	15	8.24	45.05	44.10	2.11%
68	4	0.75	2	4	200	15	8.01	46.59	46.70	0.25%
69	4	0.75	1	4	200	15	8.35	44.35	44.08	0.62%
70	4	0.75	2	4	200	15	7.86	47.57	46.60	2.06%
71	4	0.75	1	4	200	15	8.32	44.56	43.55	2.26%
72	4	0.75	2	4	200	15	7.80	48.03	46.54	3.10%
73	4	0.75	1	4	200	15	8.31	44.59	43.68	2.04%

74	4	0.75	2	4	200	15	7.80	47.99	47.55	0.91%
----	---	------	---	---	-----	----	------	-------	-------	-------

Se definió, para el estudio de aceptación de las predicciones, como criterio de aceptación (CA): $CA \leq 20\%$ en el porcentaje de error (%E) entre las predicciones y el resultado experimental (% Remediación de Cu, **Tabla 11**). Dicho criterio es adoptado del Método estándar en el capítulo de control de calidad en la determinación de metales en aguas (SMEWW 3020B.2h, 23rd) (SMEWW, 2017). Esta adopción es realizada como complemento al estudio hecho, pues no hay un criterio de aceptación definido para el porcentaje de error en predicciones utilizando el modelo de RNA.

La **Tabla 12** extrae aquellos eventos en los cuales el porcentaje de error de la predicción del %Remoción, sobre el esperado, (%E) supera el criterio de aceptación: $\%E \geq 20\%$.

Tabla 12: Errores máximos determinados comparando el %Rpredictivo vs %Rexperimental, para la RNA de la **Tabla 11**.

N°	pH	masa bp (g)	conc GLA (g/100mL)	tiempo (min)	agitación (rpm)	conc Cu inicial (mg/L)	Conc Final	% Remoción	%R_pred	%E
19	3	1	1	3	300	10	3.30	67.02	80.99	20.86%
25	3	0.5	1	5	300	10	6.48	35.20	42.62	21.09%
63	3	1	2	5	300	20	13.68	31.60	38.54	21.94%

Se ve que hay 3 valores que exceden el 20% de error. Como el modelo de la RNA (6 – 16 – 1) es de predicción a un 95% de confianza, se espera que haya máximo 5% de la cantidad de datos en total, que sería 4 datos ($74 \times 5\% = 3.7$), y como se ha dicho que hay exactamente 3 datos que están fuera de especificación (error mayor al 20%), se ve que el modelo cumple con la limitante de la significancia ($\alpha = 0.05$), demostrando que el proceso está dentro de control estadístico.

Cabe señalar que el modelo predictorio a cumplido, llegando a la baya mínima (el máximo de errores soportados). Para mejorar estos resultados se requiere claramente de

mayor cantidad de datos de patrones para el entrenamiento, logrando así un modelo más ajustado y teniendo menor índice de error.

Se realiza el análisis de regresión (**Gráfico 24**) para los datos predictivos (%R Predictivo) respecto al experimental (%R Experimental).

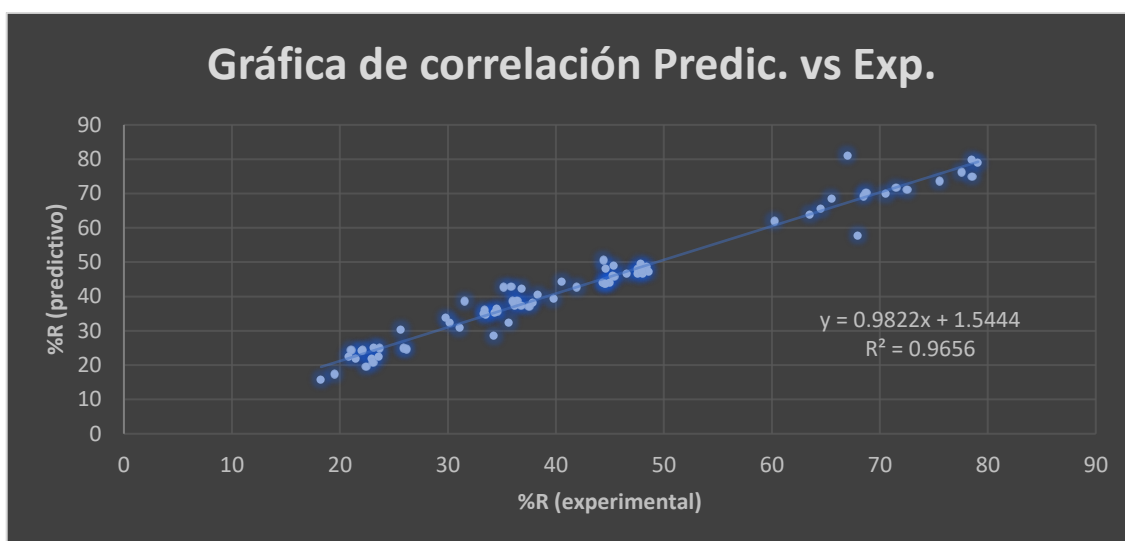


Gráfico 24: Gráfica de dispersión de los valores de %R predictivo vs experimental

Análisis de Regresión: Correlación y ANOVA de regresión. Los resultados se presentan a continuación (resultados obtenidos con Minitab ® 17):

Análisis de regresión: %R_pred vs. %Remoción

Análisis de Varianza

Fuente	GL	SC Ajust.	MC Ajust.	Valor F	Valor p
Regresión	1	20574.0	20574.0	2022.07	0.000
%Remoción	1	20574.0	20574.0	2022.07	0.000
Error	72	732.6	10.2		
Total	73	21306.6			

Resumen del modelo

S	R-cuad.	R-cuad. (ajustado)	R-cuad. (pred)
3.18979	96.56%	96.51%	96.33%

Coefficientes

Término	Coef	EE del coef.	Valor T	Valor p	VIF
Constante	1.54	1.00	1.54	0.129	
%Remoción	0.9822	0.0218	44.97	0.000	1.00

Ecuación de regresión

$$\%R_{\text{pred}} = 1.54 + 0.9822 \%Remoción$$

Del análisis de regresión se tiene:

- p-valor = 0.000 < 0.05, entonces si hay evidencia de una correlación entre las variables R% (Predictivo) vs R% (Experimental).
- Coeficiente de determinación para el modelo lineal: $R^2 = 96.56\%$, luego el coeficiente de correlación $r = 0.982649$, por lo que se concluye que el modelo lineal es adecuado para la descripción de la correlación entre las variables R% (Predictivo) vs R% (Experimental), y que el modelo lineal describe adecuadamente el comportamiento de correlación con un 96.56% de correspondencia.

La evaluación Robustez de la correlación, hecha por *Microsoft Excel Professional Plus*

2016:

<i>Estadísticas de la regresión</i>	
Coefficiente de correlación múltiple	0.98265821
Coefficiente de determinación R ²	0.96561716
R ² ajustado	0.96513963
Error típico	3.18978604
Observaciones	74

ANÁLISIS DE VARIANZA					
	<i>Grados de libertad</i>	<i>Suma de cuadrados</i>	<i>Promedio de los cuadrados</i>	<i>F</i>	<i>Valor crítico de F</i>
Regresión	1	20574.0076	20574.0076	2022.06816	1.9385E-54
Residuos	72	732.580919	10.174735		
Total	73	21306.5886			

	<i>Coefficientes</i>	<i>Error típico</i>	<i>Estadístico t</i>	<i>Probabilidad</i>	<i>Inferior 95%</i>	<i>Superior 95%</i>
Intercepción	1.54439681	1.00498714	1.53673291	0.12874192	-0.4590084	3.54780206
Variable X 1	0.9822475	0.02184354	44.9674122	1.9385E-54	0.93870319	1.0257918

Ecuación de regresión: correlación %R predictivo vs %R experimental

$$y = 0.9822x - 1.5444$$

Datos de robustez:

Pendiente: IC(95%) = [0.9387 – 1.0258]

Intercepto: IC(95%) = [-0.4590 – 3.5478]

Estos datos son de control de calidad del modelo predictivo, el cual servirá de simiento ante la aplicación futura de remoción de cobre en matrices de aguas de diferentes fuentes (p.e. residual, superficial, de efluentes industriales, salinas, etc.). Esto se debe pues las pruebas son en matriz sintética (contaminación en agua pura), sin embargo ante la pruebas y el entrenamiento al modelo de RNA (6 – 16 – 1), se deberá controlar que los datos de regresión estén inmersos en los datos de pendiente e intercepto de la matriz

sintética (libre de interferentes) en un intervalo de confianza del 95%. Esto según el criterio de esperar que la matriz sintética (libre de interferentes) dé el comportamiento teórico – ideal, el cual también dará los controles para el aseguramiento que las otras redes neuronales en matrices reales de agua puedan ser aceptadas en este criterio o desechadas por este mismo. Esto se presenta en el **Gráfico 25** donde se ven los comportamientos de la curva en un intervalo de confianza al 95%.

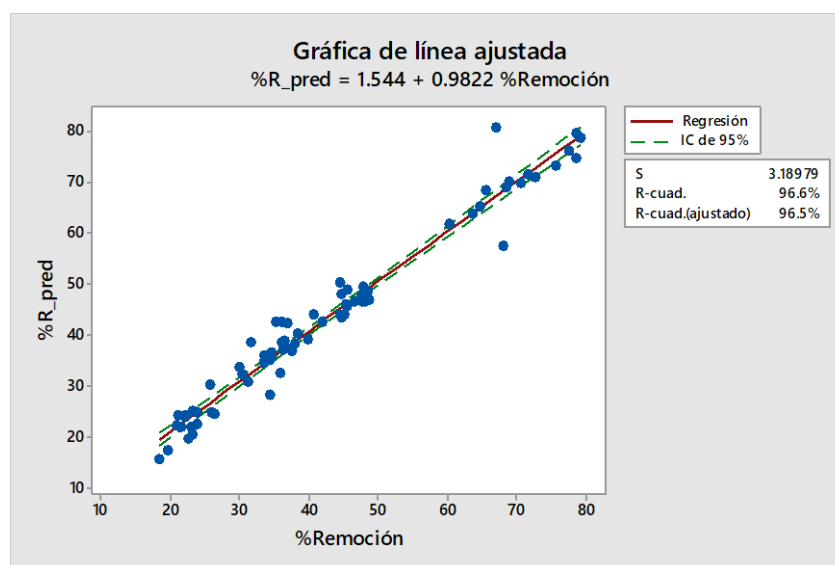


Gráfico 25: Gráfico QA/QC de la regresión %R predictiva vs %R experimental dada por la matriz sintética (agua desionizada). Gráfica hecha por Minitab ® 17.

Y el respectivo análisis residual de la regresión entre el %R predictiva vs %R experimental, generada por el Minitab ® 17 (**Gráfico 26**):

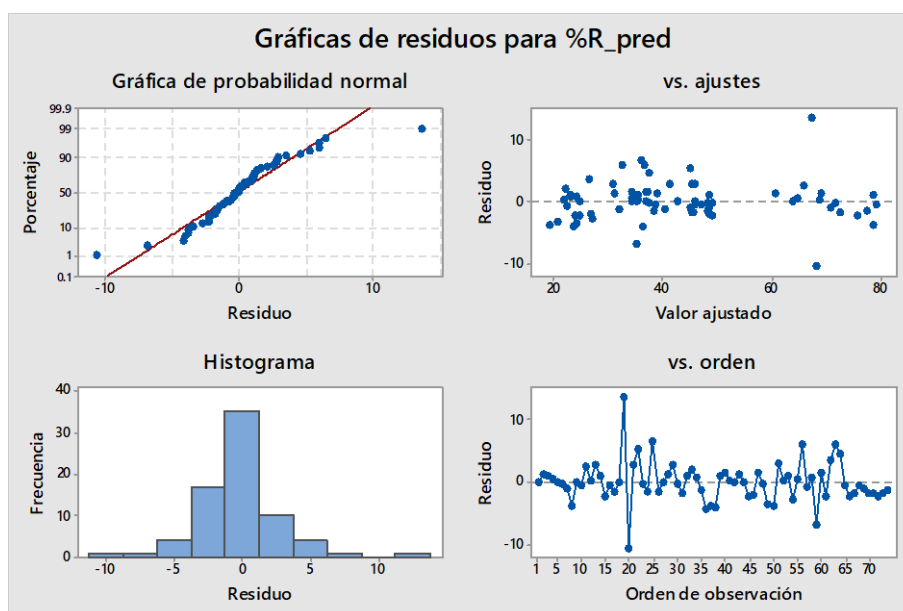


Gráfico 26: Análisis residual de la regresión %R predictiva vs %R experimental dada por la matriz sintética (agua desionizada). Gráfica hecha por Minitab ® 17.

Validación de la RNA:

Para la validación del modelo, se ha evaluado la exactitud de la respuesta de la RNA, ya entrenada, con eventos indicados en el DOE siguiente (combinaciones que antes no fueron objeto de estudio en la etapa de aprendizaje), y comparando con lo que experimentalmente se obtiene.

Tabla 13: Resultados obtenidos de la RNA previamente entrenada, con 16 capas ocultas y con los parámetros fijados para la validación.

N°	pH	masa bp (g)	conc GLA (g/100mL)	tiempo (min)	agitación (rpm)	conc Cu inicial (mg/L)	Conc Final	% Remoción	%R_pred	%E
1	4	0.5	2	5	300	10	5.33	46.68	46.73	0.10%
2	4	1	2	5	300	20	13.52	32.39	34.06	5.17%
3	7	0.5	2	5	300	10	5.35	46.48	51.32	10.42%
4	7	1	2	5	300	20	11.67	41.66	47.90	14.98%
5	5	0.5	2	5	300	10	5.34	46.64	48.26	3.47%
6	5	1	2	5	300	20	12.60	36.99	36.68	0.85%
7	2	0.5	2	5	300	10	5.62	43.76	43.67	0.21%
8	2	1	2	5	300	20	11.84	40.81	32.84	19.53%
9	7	0.25	2	5	300	10	6.14	38.65	34.79	9.99%

En la **Tabla 13** se observa que los resultados de predicción para los eventos de validación, todos ellos están por debajo del valor máximo permitido en desviación (20%), aun cuando en algunos eventos el error está casi en el límite máximo permitido. Esto se debe a que los eventos simulados en la RNA se encuentran fuera del alcance de la matriz de diseño de experimentos.

Estos datos son ingresados nuevamente a red neuronal como *patrones de aprendizaje* (eventos y % de remoción experimental), teniendo así un total de 83 registros (74 de la etapa de aprendizaje, y 9 de la etapa de validación), los cuales por predicción del modelo de sus respectivos eventos, se obtiene como sigue (**Tabla 14**):

Tabla 14: Resultados obtenidos de la RNA previamente entrenada con la matriz expandida, con 16 capas ocultas.

N°	pH	masa bp (g)	conc GLA (g/100mL)	tiempo (min)	agitación (rpm)	conc Cu inicial (mg/L)	Conc Final	% Remoción	%R_pred	%E
1	3	0.5	1	3	100	10	6.38	36.21	36.91	1.93%
2	5	0.5	1	3	100	10	6.17	38.34	39.04	1.83%
3	3	1	1	3	100	10	3.97	60.26	60.96	1.16%
4	5	1	1	3	100	10	3.54	64.58	65.29	1.08%
5	3	0.5	2	3	100	10	5.47	45.26	45.65	0.86%

6	5	0.5	2	3	100	10	5.24	47.59	47.98	0.82%
7	3	1	2	3	100	10	2.94	70.57	70.97	0.55%
8	5	1	2	3	100	10	2.14	78.57	78.97	0.50%
9	3	0.5	1	5	100	10	6.65	33.54	34.24	2.09%
10	5	0.5	1	5	100	10	6.21	37.85	38.55	1.85%
11	3	1	1	5	100	10	3.45	65.55	66.25	1.07%
12	5	1	1	5	100	10	3.15	68.50	69.20	1.02%
13	3	0.5	2	5	100	10	5.54	44.59	44.98	0.88%
14	5	0.5	2	5	100	10	5.21	47.87	48.26	0.82%
15	3	1	2	5	100	10	2.45	75.54	75.93	0.52%
16	5	1	2	5	100	10	2.09	79.10	79.50	0.49%
17	3	0.5	1	3	300	10	6.25	37.55	38.25	1.87%
18	5	0.5	1	3	300	10	5.8	41.95	42.65	1.67%
19	3	1	1	3	300	10	3.3	67.02	67.72	1.05%
20	5	1	1	3	300	10	3.2	67.99	68.69	1.03%
21	3	0.5	2	3	300	10	5.94	40.59	40.98	0.96%
22	5	0.5	2	3	300	10	5.56	44.41	44.80	0.88%
23	3	1	2	3	300	10	2.85	71.54	71.93	0.55%
24	5	1	2	3	300	10	2.24	77.59	77.98	0.50%
25	3	0.5	1	5	300	10	6.48	35.20	35.90	1.99%
26	5	0.5	1	5	300	10	6.02	39.81	40.51	1.76%
27	3	1	1	5	300	10	3.65	63.54	64.24	1.10%
28	5	1	1	5	300	10	3.12	68.76	69.46	1.02%
29	3	0.5	2	5	300	10	5.47	45.35	45.33	0.05%
30	5	0.5	2	5	300	10	5.16	48.41	47.81	1.24%
31	3	1	2	5	300	10	2.75	72.54	74.20	2.28%
32	5	1	2	5	300	10	2.15	78.54	76.41	2.72%
33	3	0.5	1	3	100	20	15.79	21.05	21.75	3.33%
34	5	0.5	1	3	100	20	15.37	23.16	23.86	3.03%
35	3	1	1	3	100	20	13.78	31.09	31.79	2.25%
36	5	1	1	3	100	20	12.87	35.67	36.37	1.96%
37	3	0.5	2	3	100	20	16.35	18.26	18.65	2.14%
38	5	0.5	2	3	100	20	15.51	22.44	22.83	1.74%
39	3	1	2	3	100	20	13.09	34.55	34.94	1.13%
40	5	1	2	3	100	20	12.71	36.44	36.83	1.07%
41	3	0.5	1	5	100	20	15.84	20.81	21.51	3.37%
42	5	0.5	1	5	100	20	15.26	23.69	24.39	2.96%
43	3	1	1	5	100	20	13.96	30.18	30.88	2.32%
44	5	1	1	5	100	20	13.13	34.34	35.04	2.04%
45	3	0.5	2	5	100	20	15.41	22.93	23.32	1.71%

46	5	0.5	2	5	100	20	14.82	25.92	26.32	1.51%
47	3	1	2	5	100	20	13.32	33.41	33.80	1.17%
48	5	1	2	5	100	20	12.64	36.79	37.18	1.06%
49	3	0.5	1	3	300	20	16.09	19.53	20.23	3.59%
50	5	0.5	1	3	300	20	15.38	23.12	23.82	3.03%
51	3	1	1	3	300	20	14.03	29.84	30.54	2.35%
52	5	1	1	3	300	20	13.08	34.60	35.30	2.02%
53	3	0.5	2	3	300	20	15.58	22.11	22.50	1.77%
54	5	0.5	2	3	300	20	14.76	26.20	26.59	1.49%
55	3	1	2	3	300	20	13.33	33.36	33.75	1.17%
56	5	1	2	3	300	20	12.83	35.87	36.26	1.09%
57	3	0.5	1	5	300	20	15.7	21.48	22.18	3.26%
58	5	0.5	1	5	300	20	15.59	22.04	22.74	3.18%
59	3	1	1	5	300	20	13.14	34.29	34.99	2.04%
60	5	1	1	5	300	20	12.79	36.06	36.76	1.94%
61	3	0.5	2	5	300	20	15.27	23.64	24.03	1.66%
62	5	0.5	2	5	300	20	14.87	25.65	26.04	1.53%
63	3	1	2	5	300	20	13.68	31.60	35.81	13.31%
64	5	1	2	5	300	20	12.63	36.83	37.36	1.45%
65	4	0.75	1	4	200	15	8.18	45.43	40.31	11.27%
66	4	0.75	2	4	200	15	7.71	48.59	45.25	6.88%
67	4	0.75	1	4	200	15	8.24	45.05	40.31	10.51%
68	4	0.75	2	4	200	15	8.01	46.59	45.25	2.88%
69	4	0.75	1	4	200	15	8.35	44.35	40.31	9.11%
70	4	0.75	2	4	200	15	7.86	47.57	45.25	4.89%
71	4	0.75	1	4	200	15	8.32	44.56	40.31	9.53%
72	4	0.75	2	4	200	15	7.8	48.03	45.25	5.79%
73	4	0.75	1	4	200	15	8.31	44.59	40.31	9.59%
74	4	0.75	2	4	200	15	7.8	47.99	45.25	5.71%
Validación										
75	4	0.5	2	5	300	10	5.33	46.68	46.57	0.24%
76	4	1	2	5	300	20	13.52	32.39	36.59	12.95%
77	7	0.5	2	5	300	10	5.35	46.48	50.29	8.20%
78	7	1	2	5	300	20	11.67	41.66	38.91	6.60%
79	5	0.5	2	5	300	10	5.34	46.64	47.81	2.51%
80	5	1	2	5	300	20	12.6	36.99	37.36	1.00%
81	2	0.5	2	5	300	10	5.62	43.76	44.09	0.75%
82	2	1	2	5	300	20	11.84	40.81	35.04	14.15%
83	7	0.25	2	5	300	10	6.14	38.65	36.12	6.53%

Se observa que en todos los patrones (incluyendo los de la validación), su error de predicción es menor de 20% (error máximo 14.15%), por lo que se puede evidenciar que los resultados de predicción se han ajustado mejor al entrenar a la RNA con los patrones de validación, que por su selección han sido valores de tendencia distinta a la de los obtenidos en el diseño de experimentos. Esto porque el algoritmo se ajusta no solo a la tendencia del diseño factorial con 5 puntos centrales, sino a otros valores que escapan de la matriz de diseño, emulando mejor el proceso de sorción.

Se procede a hacer el análisis de correlación (**Gráfico 27**) entre los resultados de la predicción, con respecto a los resultados experimentales:

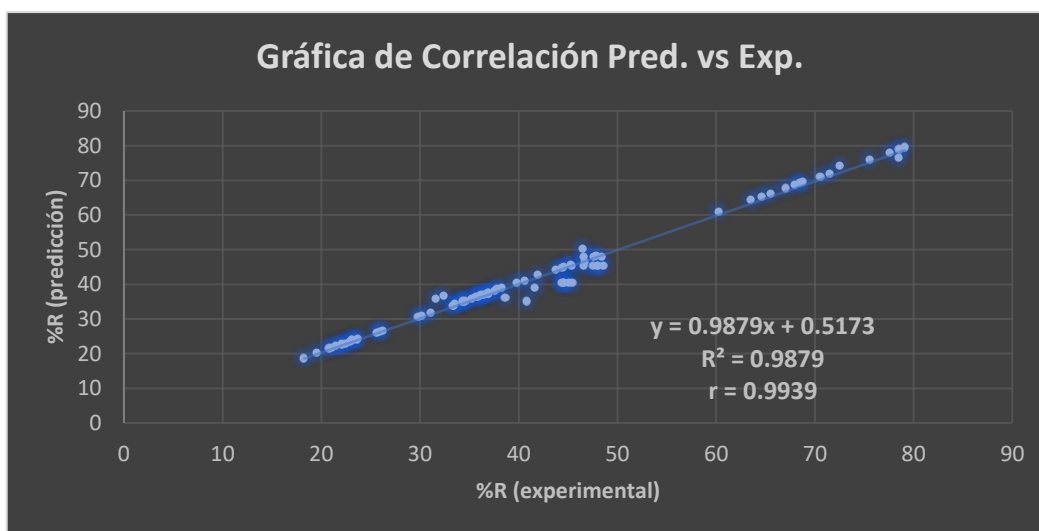


Gráfico 27: Gráfica de dispersión de los valores de %R predictivo vs experimental en la matriz de experimento expandida.

Se observa que el coeficiente de regresión de Pearson $r = 0.9939$ y el coeficiente de determinación de la regresión $R^2 = 0.9879$ han aumentado respecto a la simulación del DoE factorial ($r = 0.9656$, $R^2 = 0.9826$).

Realizando el análisis de regresión en Minitab ® 17:

Análisis de regresión: %R-pred vs. %R-exp					
Análisis de Varianza					
Fuente	GL	SC Ajust.	MC Ajust.	Valor F	Valor p
Regresión	1	21011.7	21011.7	6594.85	0.000
%R-exp	1	21011.7	21011.7	6594.85	0.000
Error	81	258.1	3.2		
Total	82	21269.8			
Resumen del modelo					
		R-cuad.	R-cuad.		
S	R-cuad.	(ajustado)	(pred)		
1.78496	98.79%	98.77%	98.75%		
Coeficientes					
Término	Coef	EE del coef.	Valor T	Valor p	VIF
Constante	0.517	0.554	0.93	0.354	
%R-exp	0.9879	0.0122	81.21	0.000	1.00
Ecuación de regresión					
%R-pred = 0.517 + 0.9879 %R-exp					

Del análisis de regresión se tiene:

- p-valor = 0.000 < 0.05, entonces si hay evidencia de una correlación entre las variables R% (Predictivo) vs R% (Experimental).
- Coeficiente de determinación para el modelo lineal: $R^2 = 98.79\%$, luego el coeficiente de correlación $r = 0.9939$, por lo que se concluye que el modelo lineal es adecuado para la descripción de la correlación entre las variables R% (Predictivo) vs R% (Experimental), y que el modelo lineal describe adecuadamente el comportamiento de correlación con un 98.79% de correspondencia.

La evaluación Robustez de la correlación, hecha por *Microsoft Excel Professional Plus 2016*:

<i>Regression Statistics</i>	
Multiple R	0.99391485
R Square	0.98786672
Adjusted R Square	0.98771693
Standard Error	1.78495949
Observations	83

<i>ANOVA</i>					
	<i>df</i>	<i>SS</i>	<i>MS</i>	<i>F</i>	<i>Significance F</i>
Regression	1	21011.7323	21011.7323	6594.85315	2.239E-79
Residual	81	258.072512	3.18608039		
Total	82	21269.8048			

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>	<i>Upper 95%</i>
Intercept	0.51726736	0.55437587	0.93306254	0.35356073	-0.5857667	1.62030144
X Variable 1	0.98786672	0.01216454	81.2087012	2.239E-79	0.9636631	1.01207034

Ecuación de regresión: correlación %Rpredictivo vs %Rexperimental

$$y = 0.9879x - 0.5173$$

Datos de robustez:

Pendiente: IC(95%) = [0.9637 – 1.0121]

Intercepto: IC(95%) = [-0.5858 – 1.6203]

Se observa que los resultados de la regresión contienen al intercepto y $(x = 0) = 0$ y la pendiente $m = 1$, demostrando así que la regresión es correcta para la predicción.

Se presenta los intervalos de confianza al 95% de la regresión en el **Gráfico 28**:

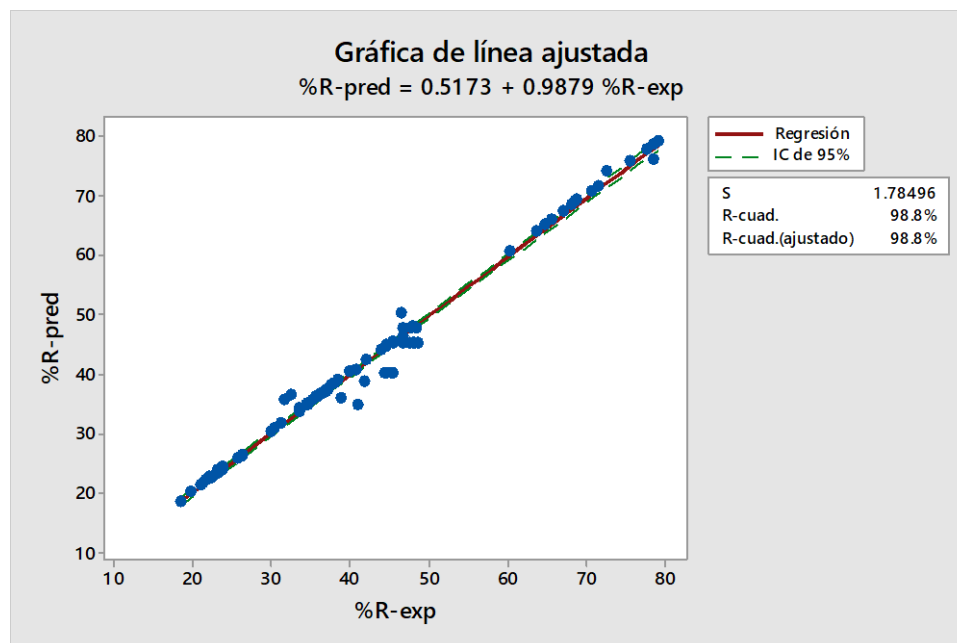


Gráfico 28: Gráfico QA/QC de la regresión %R predictiva vs %R experimental dada por la matriz expandida. Gráfica hecha por Minitab ® 17.

Se evidencia que los resultados ahora tienen un I.C. (95%) más estrecho que en el caso que se consideraba la matriz del DOE, esto es porque como se ha evidenciado en la **Tabla 14**, los resultados de predicción de la RNA han disminuido su sesgo respecto al valor de referencia experimental.

Y el respectivo análisis residual de la regresión entre el %R predictiva vs %R experimental, generada por el Minitab ® 17 (**Gráfico 29**):

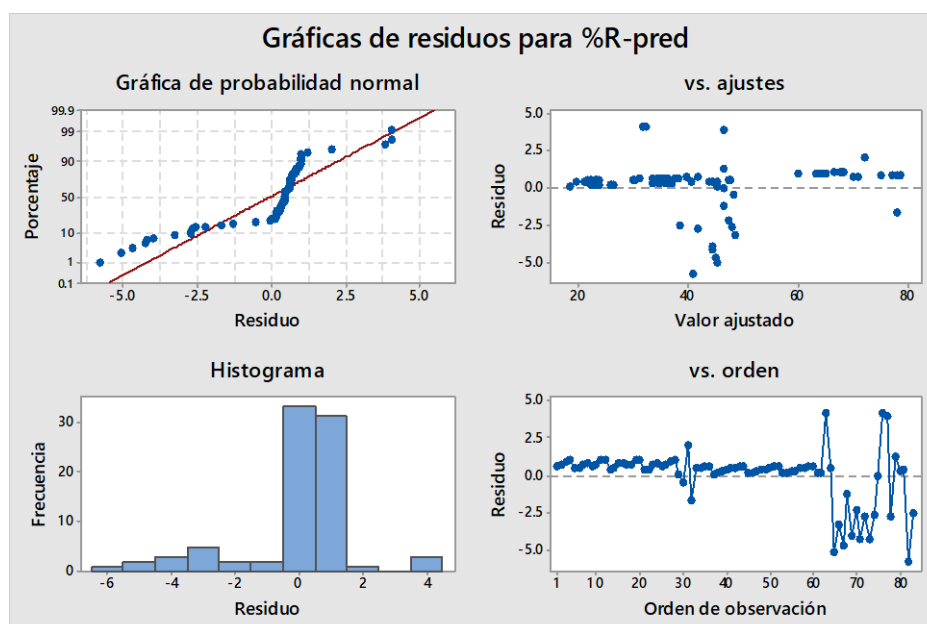


Gráfico 29: Análisis residual de la regresión %R predictiva vs %R experimental dada por la matriz expandida. Gráfica hecha por Minitab ® 17.

Finalmente, se elabora una tabla de comparación de los resultados de validación entre la RNA entrenada con la matriz del DOE y la RNA entrenada con la matriz expandida (**Tabla 15**):

Tabla 15: Comparación de los resultados obtenidos con la RNA entrenada con la matriz del DOE y la entrenada con la matriz expandida (incluyendo patrones de validación).

pH	masa bp (g)	conc GLA (g/100mL)	tiempo (min)	agitación (rpm)	conc Cu inicial (mg/L)	% Remoción	%R - RNA i	%E - RNAi	%R - RNA f	%E - RNAf	Racha %
4	0.5	2	5	300	10	46.68	46.73	0.10%	46.57	0.24%	-0.13
4	1	2	5	300	20	32.39	34.06	5.17%	36.59	12.95%	-7.79
7	0.5	2	5	300	10	46.48	51.32	10.42%	50.29	8.20%	2.2
7	1	2	5	300	20	41.66	47.92	14.98%	38.91	6.60%	8.39
5	0.5	2	5	300	10	46.64	48.26	3.47%	47.81	2.51%	0.97
5	1	2	5	300	20	36.99	36.68	0.85%	37.36	1.00%	-0.16
2	0.5	2	5	300	10	43.76	43.67	0.21%	44.09	0.75%	-0.53
2	1	2	5	300	20	40.81	32.84	19.53%	35.04	14.15%	5.38
7	0.25	2	5	300	10	38.65	34.79	9.99%	36.13	6.53%	3.46
							Promedio	7.19%	Promedio	5.88%	11.80%

Se puede visualizar en la **Tabla 15** que ha mejorado en una gran medida la predicción del %Remoción en la RNA entrenada con la matriz expandida (DOE inicial y los patrones de validación), con respecto a la RNA entrenada con la matriz del DOE. Este error de predicción de la RNA entrenada con la matriz expandida (que es en promedio 5.88%) mejorará en cuanto se entrene a la RNA con otros patrones distintos a la matriz del DOE.

Finalmente se ha visto que al término de la realización del experimento de los eventos descritos en la matriz expandida, muchas de las perlas de quitosano encrustadas con glutaraldehído se destruían, producto de las revoluciones de giro y las colisiones entre perlas, haciendo así que disminuya la eficacia de la remoción, siendo ello una de las razones por la cual se obtuvo, en algunos experimentos, un alto error de predicción por la comparación con resultados de remoción experimental con eficacia no esperada (perlas desechas o partidas).

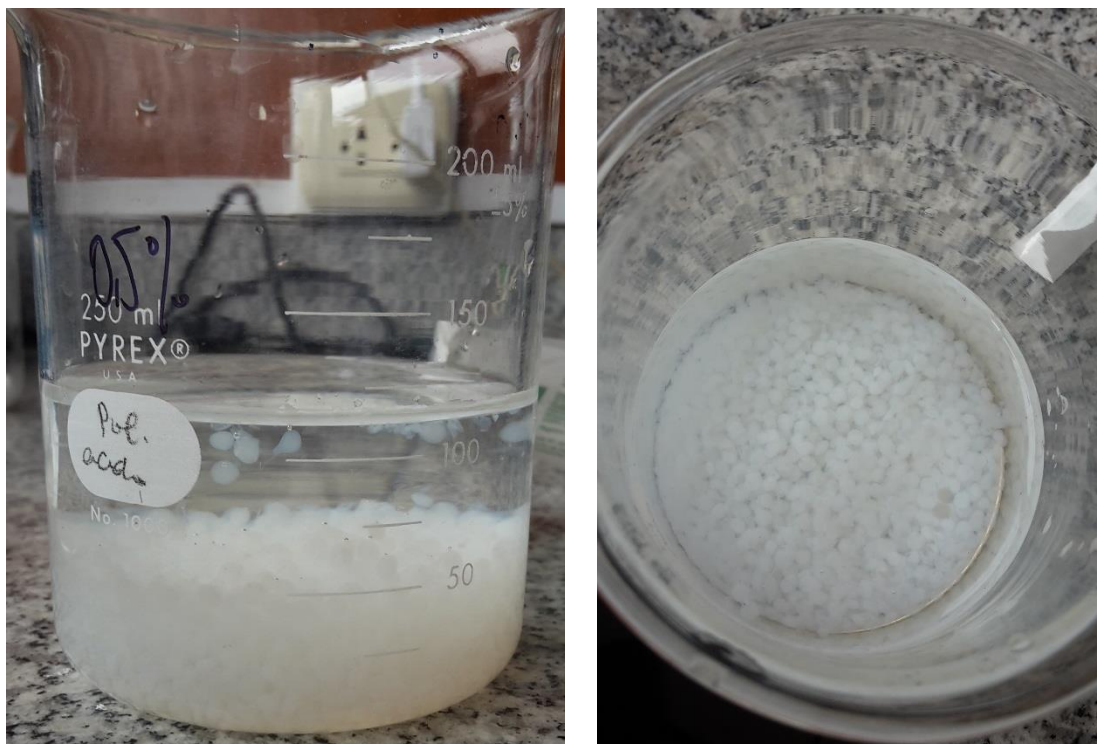


Gráfico 30: Perlas de quitosano entrecruzadas con glutaraldehído.

Se presenta los datos de optimización del experimento, buscando la máxima remoción en 20ppm de Cu en un volumen de 100 mL de agua, fijando el parámetro de tiempo (5 min) y frecuencia de agitación (300rpm) en los niveles básico, neutro y ácido. Se presenta un resumen de escenarios de optimización en la **Tabla 16**:

Tabla 16: Resultados de optimización (%R al 100% de Cu) en función de la masa de quitosano / GLA, en nivel ácido, neutro y básico.

pH	masa bp (g)	conc GLA (g/100mL)	tiempo (min)	agitación (rpm)	conc Cu inicial (mg/L)	% Remoción
4	1.44461	2	5	300	10	100
4	3.67140	2	5	300	20	100
7	3.71064	2	5	300	20	100
10	3.75440	2	5	300	20	100

Finalmente, se hace la comparación de la predicción de la RNA, con el diseño factorial completo 2⁶ (+5PC) y una regresión lineal múltiple, considerando los 9 puntos de validación, presentándose en la **Tabla 17**:

Tabla 17: Comparación de las predicciones del RNA, DOE 2⁶ (+5PC) y regresión lineal múltiple.

Nro	% Remoción	RNA		DOE 2 ⁶ (+5PC)		Regresión Lineal Múltiple	
		% R (predicción)	% Error	% R (predicción)	% Error	% R (predicción)	% Error
1	36.21	36.91	-1.93%	36.71	0.55%	42.18	-14.27%
2	38.34	39.04	-1.83%	36.62	6.20%	44.59	-14.23%
3	60.26	60.96	-1.16%	36.57	40.01%	62.54	-2.59%
4	64.58	65.29	-1.08%	36.68	43.81%	64.95	0.52%
5	45.26	45.65	-0.86%	36.15	20.80%	46.43	-1.71%
6	47.59	47.98	-0.82%	35.63	25.74%	48.84	-1.80%
7	70.57	70.97	-0.55%	36.00	49.27%	66.79	5.89%
8	78.57	78.97	-0.50%	35.62	54.89%	69.20	12.37%
9	33.54	34.24	-2.09%	36.16	-5.58%	42.27	-23.44%
10	37.85	38.55	-1.85%	35.63	7.59%	44.68	-15.90%
11	65.55	66.25	-1.07%	36.00	45.66%	62.63	5.47%
12	68.50	69.20	-1.02%	35.62	48.53%	65.04	6.01%
13	44.59	44.98	-0.88%	34.01	24.39%	46.52	-3.43%

14	47.87	48.26	-0.82%	32.98	31.66%	48.93	-1.39%
15	75.54	75.93	-0.52%	33.65	55.69%	66.88	11.93%
16	79.10	79.50	-0.49%	32.71	58.85%	69.29	12.84%
17	37.55	38.25	-1.87%	60.77	-58.88%	43.32	-13.27%
18	41.95	42.65	-1.67%	61.36	-43.85%	45.74	-7.22%
19	67.02	67.72	-1.05%	60.39	10.82%	63.68	5.96%
20	67.99	68.69	-1.03%	61.40	10.61%	66.09	3.78%
21	40.59	40.98	-0.96%	62.44	-52.38%	47.57	-16.09%
22	44.41	44.80	-0.88%	62.28	-39.02%	49.99	-11.58%
23	71.54	71.93	-0.55%	62.24	13.48%	67.93	5.57%
24	77.59	77.98	-0.50%	62.42	19.96%	70.34	9.80%
25	35.20	35.90	-1.99%	62.44	-73.93%	43.413	-20.92%
26	39.81	40.51	-1.76%	62.28	-53.74%	45.83	-13.12%
27	63.54	64.24	-1.10%	62.24	3.12%	63.77	0.74%
28	68.76	69.46	-1.02%	62.42	10.14%	66.18	4.72%
29	45.35	45.33	0.05%	61.48	-35.63%	47.66	-5.15%
30	48.41	47.81	1.24%	60.56	-26.66%	50.08	-4.74%
31	72.54	74.20	-2.28%	61.27	17.42%	68.02	8.33%
32	78.54	76.41	2.72%	60.64	20.63%	70.43	7.82%
33	21.05	21.75	-3.33%	36.71	-68.78%	14.45	33.56%
34	23.16	23.86	-3.03%	36.62	-53.48%	16.86	29.32%
35	31.09	31.79	-2.25%	36.57	-15.06%	34.81	-9.50%
36	35.67	36.37	-1.96%	36.68	-0.86%	37.22	-2.33%
37	18.26	18.65	-2.14%	36.15	-93.88%	18.70	-0.28%
38	22.44	22.83	-1.74%	35.63	-56.03%	21.11	7.54%
39	34.55	34.94	-1.13%	36.00	-3.03%	39.06	-11.78%
40	36.44	36.83	-1.07%	35.62	3.28%	41.47	-12.60%
41	20.81	21.51	-3.37%	36.16	-68.09%	14.54	32.40%
42	23.69	24.39	-2.96%	35.63	-46.08%	16.95	30.49%
43	30.18	30.88	-2.32%	36.00	-16.56%	34.90	-13.00%
44	34.34	35.04	-2.04%	35.62	-1.65%	37.31	-6.48%
45	22.93	23.32	-1.71%	34.01	-45.82%	18.79	19.43%
46	25.92	26.32	-1.51%	32.98	-25.32%	21.20	19.43%
47	33.41	33.80	-1.17%	33.65	0.45%	39.15	-15.81%
48	36.79	37.18	-1.06%	32.71	12.03%	41.56	-11.76%
49	19.53	20.23	-3.59%	60.77	-200.39%	15.59	22.92%
50	23.12	23.82	-3.03%	61.36	-157.61%	18.01	24.41%
51	29.84	30.54	-2.35%	60.39	-97.74%	35.95	-17.71%
52	34.60	35.30	-2.02%	61.40	-73.93%	38.36	-8.67%
53	22.11	22.50	-1.77%	62.44	-177.49%	19.84	11.82%
54	26.20	26.59	-1.49%	62.28	-134.19%	22.26	16.31%
55	33.36	33.75	-1.17%	62.24	-84.42%	40.20	-19.12%
56	35.87	36.26	-1.09%	62.42	-72.12%	42.61	-17.50%
57	21.48	22.18	-3.26%	62.44	-181.51%	15.68	29.30%

58	22.04	22.74	-3.18%	62.28	-173.83%	18.10	20.44%
59	34.29	34.99	-2.04%	62.24	-77.85%	36.04	-2.99%
60	36.06	36.76	-1.94%	62.42	-69.79%	38.45	-4.60%
61	23.64	24.03	-1.66%	61.48	-155.79%	19.93	17.07%
62	25.65	26.04	-1.53%	60.56	-132.52%	22.35	14.20%
63	31.60	35.81	-13.31%	61.27	-71.10%	40.29	-12.50%
64	36.83	37.36	-1.45%	60.64	-62.31%	42.70	-14.29%
65	45.43	40.31	11.27%	29.30	27.33%	40.32	-0.01%
66	48.59	45.25	6.88%	27.43	39.37%	44.57	1.51%
67	45.05	40.31	10.51%	29.30	27.33%	40.32	-0.01%
68	46.59	45.25	2.88%	27.43	39.37%	44.57	1.51%
69	44.35	40.31	9.11%	29.30	27.33%	40.32	-0.01%
70	47.57	45.25	4.89%	27.43	39.37%	44.57	1.51%
71	44.56	40.31	9.53%	29.30	27.33%	40.32	-0.01%
72	48.03	45.25	5.79%	27.43	39.37%	44.57	1.51%
73	44.59	40.31	9.59%	29.30	27.33%	40.32	-0.01%
74	47.99	45.25	5.71%	27.43	39.37%	44.57	1.51%
75	46.68	46.57	0.24%	30.65	34.18%	47.73	-2.49%
76	32.39	36.59	-12.95%	30.38	16.97%	40.35	-10.29%
77	46.48	50.29	-8.20%	49.80	0.98%	51.35	-2.10%
78	41.66	38.91	6.60%	49.35	-26.84%	43.97	-13.00%
79	46.64	47.81	-2.51%	32.98	31.02%	48.93	-2.35%
80	36.99	37.36	-1.00%	32.71	12.44%	41.56	-11.23%
81	43.76	44.09	-0.75%	41.70	5.42%	45.32	-2.78%
82	40.81	35.04	14.15%	41.22	-17.65%	37.94	-8.29%
83	38.65	36.12	6.53%	56.31	-55.88%	41.17	-13.96%

Se puede observar que los errores máximos obtenidos, por modelo de predicción (*Tabla 18*):

Tabla 18: Errores máximos y error cuadrático medio de tres modelos de predicción: RNA, DOE, regresión lineal múltiple

Modelo	Error máximo	# % errores > 20%	SCM – Error promedio
RNA	14.15%	0	4.34%
DOE	-200.39%	59	65.78%
RLM	33.56%	10	13.27%

Como referencias:

- Del análisis del DOE 2⁶ (+5PC), su ecuación:

Ecuación de regresión en unidades no codificadas

$$\begin{aligned} \% \text{ Recuperación} = & 1.6 + 2.25 \text{ pH} + 93.37 \text{ masa bp (g)} + 7.28 \text{ Conc GLA (g/100mL)} \\ & + 0.61 \text{ tiempo (min)} - 0.0038 \text{ agitación (rpm)} + 0.218 \text{ Conc inicial Cu (mg/L)} \\ & + 0.321 \text{ pH*masa bp (g)} - 0.074 \text{ pH*Conc GLA (g/100mL)} - 0.274 \text{ pH*tiempo (min)} \\ & + 0.00240 \text{ pH*agitación (rpm)} - 0.0510 \text{ pH*Conc inicial Cu (mg/L)} \\ & + 1.05 \text{ masa bp (g)*Conc GLA (g/100mL)} + 0.44 \text{ masa bp (g)*tiempo (min)} \\ & - 0.0030 \text{ masa bp (g)*agitación (rpm)} \\ & - 3.638 \text{ masa bp (g)*Conc inicial Cu (mg/L)} \\ & + 0.139 \text{ Conc GLA (g/100mL)*tiempo (min)} \\ & - 0.00199 \text{ Conc GLA (g/100mL)*agitación (rpm)} \\ & - 0.3489 \text{ Conc GLA (g/100mL)*Conc inicial Cu (mg/L)} \\ & + 0.00074 \text{ tiempo (min)*agitación (rpm)} \\ & + 0.0076 \text{ tiempo (min)*Conc inicial Cu (mg/L)} \\ & - 0.000189 \text{ agitación (rpm)*Conc inicial Cu (mg/L)} + 2.979 \text{ Pt Ctrial} \end{aligned}$$

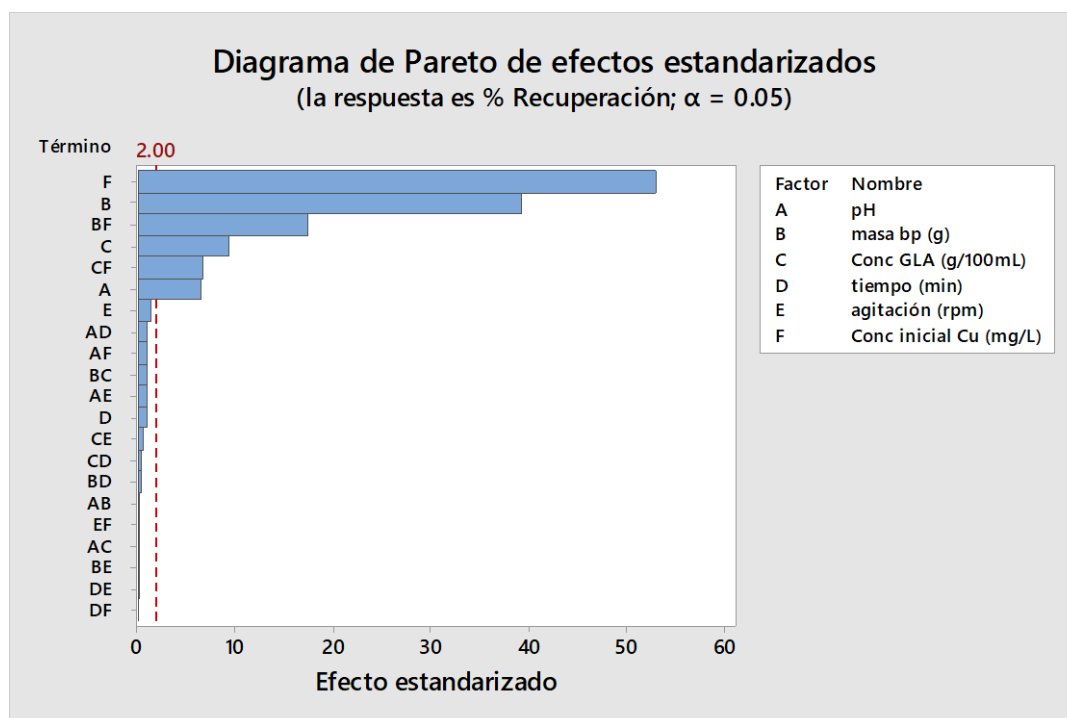


Gráfico 31: Gráfico de Pareto del análisis del DOE 2⁶ (+5PC)

Del **Gráfico 31**, se puede observar que de los 6 factores, se tienen sólo 4 factores críticos: pH, masa de BP (g), Concentración de GLA (g/100mL), Concentración inicial de Cu (mg/L), y 2 factores no significativos: tiempo de agitación (min), revoluciones de agitación (rpm).

Esto se puede ver en el **Gráfico 32** en donde se aprecia claramente cual es el efecto de cada entrada de la RNA.

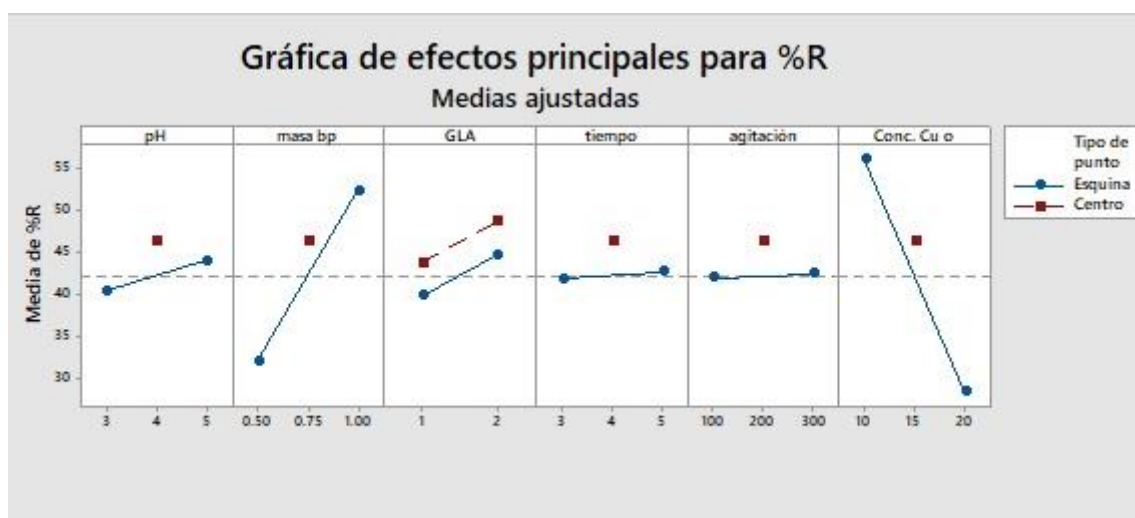


Gráfico 32: Gráfica de los efectos de los factores principales para el %Remoción

Se evidencia que tanto el tiempo de agitación como la frecuencia de agitación dan prácticamente un efecto nulo en el %Remoción del Cu(II), lo cual se confirma lo obtenido en el **Gráfico 31**. La agitación al momento del contacto entre la solución contaminada con Cu(II) y las perlas preparadas es importante, sin embargo, el cambio de las revoluciones no hace la diferencia en el fenómeno de remoción. En contra partida: los componentes que generan un efecto fuerte es: masa de biopolímero – perlas (cuyo efecto, en una variación de 0.5 a 1.0g de perlas producen una variación directa de aproximadamente 20%), y de la concentración inicial de Cu(II) (cuyo efecto, en una variación de 10 a 20 mgCu/L produce una variación inversa de aproximadamente 27%).

- Del análisis de regresión lineal múltiple, el ANOVA:

Análisis de Varianza					
Fuente	GL	SC Ajust.	MC Ajust.	Valor F	Valor p
Regresión	6	19566.4	3261.1	126.15	0.000
pH	1	143.5	143.5	5.55	0.021
masa bp (g)	1	7689.7	7689.7	297.47	0.000
Conc GLA (g/100mL)	1	361.1	361.1	13.97	0.000
tiempo (min)	1	0.1	0.1	0.01	0.941
agitación (rpm)	1	22.9	22.9	0.89	0.350
Conc inicial Cu (mg/L)	1	13771.2	13771.2	532.73	0.000
Error	76	1964.6	25.9		
Falta de ajuste	66	1960.9	29.7	79.12	0.000
Error puro	10	3.8	0.4		
Total	82	21531.0			

En el cual se puede evidenciar que tanto el tiempo y las revoluciones de agitación no se pueden modelar correctamente con el análisis de regresión, pues tienen una tendencia independiente (valor $p > 0.05$).

La ecuación de regresión respectiva, y su coeficiente de regresión, así como de determinación:

Ecuación de regresión	
% Recuperación	= 39.84 + 1.206 pH + 40.71 masa bp (g) + 4.25 Conc GLA (g/100mL) + 0.045 tiempo (min) + 0.00571 agitación (rpm) - 2.773 Conc inicial Cu (mg/L)

Resumen del modelo			
S	R-cuad.	R-cuad. (ajustado)	R-cuad. (pred)
5.08433	90.88%	90.16%	89.08%

De la estadística de la regresión lineal múltiple se puede deducir que se tiene un coeficiente de correlación entre la respuesta (%Remoción) y las 6 variables analizadas de $r = 0.9533$ ($r = \sqrt{R^2} = \sqrt{0.9088} = 0.9533$), mostrando así la fuerte correlación positiva entre las 6 variables analizadas y la respuesta (% Remoción del Cu (II)).

CONCLUSIONES

CONCLUSIONES

Con el presente trabajo de investigación se logró:

- Preparar las perlas de quitosano entrecruzadas con glutaraldehído, las cuales sirvieron para las pruebas de sorción dentro del laboratorio con agua desionizada y contaminada intencionalmente con la solución de Cu (II).
- Experimentar el fenómeno de sorción, variando 6 factores experimentales: pH de la muestra acuosa, masa de perlas de quitosano entrecruzado con glutaraldehído, concentración de glutaraldehído utilizada para el entrecruzamiento, tiempo de agitación, frecuencia de agitación, y concentración inicial de cobre. El DOE permitió identificar los factores más significativos en las pruebas: pH, masa de bp (g), concentración de GLA (g/100mL), y concentración inicial de Cu en la muestra a remover (mg/L). Además, por el gráfico de efectos de las variables principales se pudo corroborar que los factores más ponderativos en el sistema de sorción de Cu(II) por perlas de quitosano entrecruzado con glutaraldehído son la masa de perlas utilizadas y la concentración inicial del cobre, variando en 20 y 27% el porcentaje de remoción, respectivamente.
- Diseñar una red neuronal artificial multicapa (con 16 capas ocultas) con aprendizaje backpropagation capaz de poder predecir la remoción de cobre en aguas desionizadas, con un error máximo de 14.15% respecto a los valores de remoción esperado, mejorando la predicción del DOE (-200.39%) y de la regresión lineal múltiple (33.56%), y teniendo un error cuadrático medio de 4.34% para la RNA, comparada con el DOE (65.78%) y regresión lineal múltiple (13.27%).
- Realizar el análisis estadístico de la red neuronal artificial diseñada: El error de la RNA puede disminuir teniendo una data más densa de patrones, es decir, teniendo más patrones de aprendizaje para la red (en este trabajo se vio que con $n = 74$ patrones,

se tuvo un error máximo de 21.94%, y para n= 83 patrones, se tuvo un error máximo de 14.15%). Esta red neuronal artificial puede predecir el porcentaje de remoción de cobre teniendo como parámetros de ingreso el pH, masa de perlas de biopolímero quitosano / GLA, concentración molar de GLA entrecruzado en quitosano, tiempo de agitación, frecuencia de agitación y concentración inicial de cobre.

NUEVOS TRABAJOS DE INTERÉS PARA EL LECTOR

1. Verificar, y si es necesario, modificar, la predicción de la red neuronal artificial en aguas de diferentes matrices: aguas residuales (industrial), agua natural (superficial) y agua salina (agua de mar).
2. Validar este modelo en matrices reales de agua, como son: aguas residuales (industrial), agua natural (superficial) y agua salina (agua de mar).

REFERENCIAS BIBLIOGRÁFICAS

- Alsarra, I., & al, e. (2002). Molecular weight and degree of deacetylation effects on lipase-loaded chitosan bead characteristics. *Biomaterials*, 23, 3637-3644.
- Amici, E. (1995). *Sintesi di Reticolati Polimerici. Processi convenzionali e a templatato*. Dipartimento di Chimica. Roma: Università "La Sapienza".
- Barceloux, D. (1999). Copper. *Journal Clinical Toxicology*, 37, 217-230.
- Benaissa, H., & Benguella, H. (2002). Effect of anions and cations on cadmium sorption kinetic by chitin. *Proceeding of international symposium an environmental pollutions control and waste managent*, 815-822.
- Beshgetoor, & Hambidge. (1998). Clinical conditions altering copper metabolism in humans. *Am J Clin Nutr*, 67, 1017S-1021S.
- Blanco, M., & Cerdà, V. (2007). *Temas avanzados de quimiometría*. Palma: Universitat de les Illes Balears.
- Burriel, M., Conde, L., Jimeno, A., & Méndez, H. (1994). *Química Analítica Cualitativa*. Madrid: SA Paraninfo.
- Cai, W., Ye, Y., & Chem, P. (1999). *Huanjing Woman Yu Fangzhi*, 21(4), 1-4.
- Campbell, J. (2001). Lifestyle, minerals and health. *Med Hypotheses*, 57, 521-531.
- Cojocar, C., Samoila, P., & Pascariu, P. (08 de Noviembre de 2018). Chitosan-based magnetic adsorbent for removal of water-soluble anionic dye: Artificial neural network modeling and molecular docking insights. *ELSEVIER, Volumen 123*, 587-599.
- Crescenzi, V., Imbriaco, D., Lárez, C., Dentini, M., & Ciferri, A. (1995). *Macromol. Chem Phys*, 196, 2873-2880.
- De la Cruz D, C. A. (2005). Blood copper levels in Mexican users of the T380A IUD. *Contraception*, 72, 122-125.
- Díaz Sánchez, C. E. (2011). Preparación y caracterización del la sal de quitosano cuaternizado y su entrecruzamiento con glutaraldehido para la sorción del Cromo (VI) de soluciones acuosas. *Tesis para por el título profesional de licenciado en ciencias con mención en química*. Lima: Universidad Nacional de Ingeniería.
- Dolatabadi, M., Mehrabpour, M., Esfandyari, M., Alidadi, H., & Davoudi, M. (15 de Octubre de 2018). Modeling of simultaneous adsorption of dye and metal ion by sawdust from aqueous solution using of ANN and ANFIS. *ELSEVIER, volumen 181*, 72-78.

- Eddleston M, B. D. (2007). Pesticides. Medicine. *The British Journal of Psychiatry*, 35, 646-648.
- Gaetke, L., & Chow, C. (2003). Copper toxicity, oxidative stress, and antioxidant nutrients. *Toxicology*, 189, 147-163.
- Gautam, A., & Asif, H. (23 de Abril de 2018). Valorization of aquaculture waste in removal of cadmium from aqueous solution: optimization by kinetics and ANN analysis. *Springer, Applied Water Science* 8:68.
- Ghanavati Nasab, S., Semnani, A., Teimouri, A., Javaheran Yazd, M., Momeni Isfahani, T., & Habibollahi, S. (14 de Noviembre de 2018). Decolorization of crystal violet from aqueous solutions by a novel adsorbent chitosan/nanodiopside using response surface methodology and artificial neural network-genetic algorithm. *ELSEVER, Vol. 124*, 429-443.
- Greger, J. (1999). Nondigestible carbohydrates and mineral bioavailability. *Journal of Nutritional*, 129, 1434S-1435S.
- Gunay N, Y. C. (2006). A series of patients in the emergency department diagnosed with copper poisoning. *Tohoku J Exp Med*, 209, 243-248.
- Hilera, J. R., & Martínez, V. J. (2000). *Redes Neuronales Artificiales*. Madrid, España: Alfaomega.
- Hirano, S., Noishiki, Y., Kinugawa, J., Highasijima, H., & Hayashi, T. (1997). *Advances in Biomedical Polymers*. New York: C.G. Gebelein.
- Hung, I., Suzuki, M., Yamaguchi, Y., Yuan, D., Klausner, R., & Gitlin, J. (1997). Biochemical characterization of the Wilson disease protein and functional expression in the yeast *Saccharomyces cerevisiae*. *Journal of Biolo Chemistry*, 272, 21461-21466.
- IPCS. (1998). Copper. *Organización Mundial de la Salud, Programa Internacioal de Seguridad de las Sustancias Químicas*, 200.
- Iqbal, S., & Akbar, J. (27 de Diciembre de 2017). Impact of metal ionic characteristics on adsorption potential of *Ficus carica* leaves using QSPR modeling. *Journal of Environmental Science and Health*, Volumen 53, Número 4.
- Jean, C., & Holl, W. (2003). Chemical modification of chitosan and Equilibrium Study for Mercury Ion Removal. *Water Research*, 37(19), 4770-4780.
- Jia, Z., Shen, D., & Xu, W. (2001). *Carbohydr. Reserch*, 333(1), 1-6.
- Kashif Uddin, M., Ali Khan Rao, R., & Chandra Mouli, K. (15 de Septiembre de 2018). The artificial neural network and Box-Behnken design for Cu²⁺ removal by the

pottery sludge from water samples: Equilibrium, kinetic and thermodynamic studies. *ELSEVIER, Volumen 266*, 617-627.

- Kim, C., Jung, B., Choi, K., & kim, J. (1996). *Kongopo Hwahak*, 7(1), 118-128.
- Lárez, C., Canelón, F., Millán, E., & Katime, I. (2002). Synthesis and characterization of complexes between polym. *Polym. Bull.*, 48, 361-366.
- Lárez, C., Crescenzi, V., Dentini, M., & Ciferri, A. (1995). *Supramolecular Science*, 2, 141-147.
- Lasheras Zubiate, M. (2011). *Estudio de la interacción quitosano-metales pesados y su inmovilización en matrices cementicias*. Universidad de Navarra, Facultad de Ciencias.
- Linder, M. (2001). Copper and genomic stability in mammals. *Mutat Res*, 475, 141-152.
- M Thanou, A. K., Scharinghausen, T., Luessen, H., De-Boer, A., Verhoef, J., & Junginger, H. (2000). *Journal Control. Release*, 64(1-3), 15-25.
- Mahdi Niknam, S., Morteza, E., & Maryam, K. (30 de Agosto de 2018). Efficient prediction of water vapor adsorption capacity in porous metal–organic framework materials: ANN and ANFIS modeling. *Springer*, 1-10.
- Marinela, C. (Enero de 2014). Evaluación de los procesos para la obtención química de quitina y quitosano a partir de desechos de cangrejos. Escala piloto e industrial. *Revista Iberoamericana* 12(1):21.
- Mi, F., Shyu, S., Kuan, C., Lee, S., Lu, K., & Jang, S. (1999). Chitosan–Polyelectrolyte complexation for the preparation of gel beads and controlled release of anticancer drug. I. Effect of phosphorous polyelectrolyte complex and enzymatic hydrolysis of polymer. *Journal of Applied Polymer Science*, 74(7), 1868-1879.
- Ministerio del Ambiente. (07 de Junio de 2017). Decreto Supremo N° 004 - 2017 - MINAM. *El Peruano*, págs. 10-19.
- Moreno-Pérez, J., Bonilla-Petriciolet, A., Mendoza-Castillo, D., Reynel-Ávila, H., Verde-Gómez, Y., & Trejo-Valencia, R. (Agosto de 2018). Artificial neural network-based surrogate modeling of multi-component dynamic adsorption of heavy metals with a biochar. *ELSEVIER, Volumen 6*, 5389-5400.
- Muller, P., & Klomp, L. (2009). A novel copper – responsive transcription factor in mammals? *Journal Intitute Biochemical cell B*, 41, 1233-1236.
- Muzzarelli, R. (1974). *Chitin*. Italy: Perfamon Press.
- Nawaz M, M. C. (2006). Copper-induced stimulation of extracellular signal-regulated kinase in trout hepatocytes: the role of reactive oxygen species, Ca²⁺, and cell

- energetics and the impact of extracellular signal-regulated kinase signaling on apoptosis and necrosis. *Toxicol Sciences*, 92, 464-475.
- Nemery, B. (1990). Metal toxicity and the respiratory tract. *Eur Respir Journal*, 3, 202-219.
- OMS. (2003). Cooper in drinking-water. *Documento de referencia para la elaboración de las Guías de la OMS para la calidad de agua potable*, WHO/SDE/WSH/03.04/88.
- OMS. (2003). Copper in drinking-water. *Documento de referencia para la elaboración de las guías de la OMS para la calidad del agua potable*.
- Pandit, A., & Bhave, S. (1996). Present interpretation of the role of copper in Indian childhood cirrhosis. *Am Journal Clinical Nutricional*, 63, 830S-835S.
- Paradossi, G., Lisi, R., Paci, M., & Crescenzi, V. (1996). *Polym. Sci.*, A34, 3417-3425.
- Quer-Brossa, S. (1983). *Toxicología industrial*. Barcelona: España.
- Rao Karri, R., & Sahu, J. (01 de Septiembre de 2018). Process optimization and adsorption modeling using activated carbon derived from palm oil kernel shell for Zn (II) disposal from the aqueous environment using differential evolution embedded neural network. *ELSEVIER, Volumen 265*, 592-602.
- Riedel, G. (2008). *Copper. Ecotoxicol.* USA: Elsevier B.V.
- Rotilio G, C. M. (2000). Copper-dependent oxidative stress and neurodegeneration. *IUBMB Life*, 50, 309-314.
- Russell, S., & Norvig, P. (2004). *Inteligencia Artificial, un enfoque moderno* (2da ed.). Madrid, España: Pearson.
- Russell, S., & Norving, P. (2004). *Inteligencia Artificial un enfoque moderno*. Madrid: Pearson.
- Shanmugarprakash, M., Venkatachalam, S., Rajendran, K., & Pugazhendhi, A. (01 de Diciembre de 2018). Biosorptive removal of Zn(II) ions by Pongamia oil cake (*Pongamia pinnata*) in batch and fixed-bed column studies using response surface methodology and artificial neural network. *ELSEVIER, Volume 227*, 216-228.
- Sharunda D. Buchanan, R. A. (1994). *Copper in Drinking Water*. Nebraska: International Journal of Occupational and Environmental Health.
- Skjak-Braek, G., Anthonsen, T., & Sandford, P. (1989). *Chitin and Chitosan: Sources, Chemistry, Biochemistry, Physical properties and application*. New York: Elsevier.
- Slater, J., Hazen, S., & Sakthivel, S. (1993). On selecting appropriate technology for knowledge systems. *Journal of Sitems Management*, 44, 15.

- SMEWW. (2017). Standard Methods for the Examination of Water and Wastewater. En *Quality Assurance / Quality Control* (pág. 3020B.2h).
- Sushovan, S., Somnath, N., & Susmita, D. (22 de August de 2018). Application of RSM and ANN for optimization and modeling of biosorption of chromium(VI) using cyanobacterial biomass. *Indian J Chem Technol* 15, 397 - 402.
- Sutherland, C., Marcano, A., & Chitto, B. (2018). Artificial Neural Network-Genetic Algorithm Prediction of Heavy Metal Removal Using a Novel Plant-Based Biosorbent Banana Floret: Kinetic, Equilibrium, Thermodynamics and Desorption Studies. *IntechOpen*, 19.
- Tanner, M. (1998). Role of copper in Indian childhood cirrhosis. *Am Journal Clinical Nutricional*, 67, 1074S-1081S.
- Thomas, C., & Oates, P. (2003). Copper deficiency increases iron absorption in the rat. *Am Journal Physiol Gastrointest Liver Physiol*, 285, G789-795.
- Tsezos, M. (2001). Biosorption of Metals. The Experience accumulated and the outlook for technology development. *Hidrometallurgy*, 59(2-3), 241-243.
- Uriu-Adams JY, K. C. (2005). Copper, oxidative stress and human health. *Mol Aspects*, 26, 268-298.
- Vicent A, A. J.-J. (2009). Protectant activity of reduced concentration of copper sprays against *Alternaria* brown spot on “Fortune” mandarin in Spain. *European Journal of Plant Pathology*, 28, 1-6.
- Volesky, B. (1986). *Biosorption of Heavy Metals*. Boca Ratón, USA: CRC Press Inc.
- Zerounian, N., & Linder, M. (2002). Effects of copper and ceruloplasmin on iron transport in the Caco 2 cell intestinal model. *Journal Nutricional Biochemistry*, 13, 138-148.
- Zhan, Y., & Banks, C. (2005). A comparison of the properties of Polyurethane immobilized Sphagnum moss, seaweed, sunflower waste and maize for the biosorption of Cu, Pb, Zn and Ni in continuous flow packed columns. *Water Research*, 40(4), 788-798.

ANEXO

Algoritmo: redes neuronales artificiales multicapa con aprendizaje backpropagation

Explicativo:

Algoritmo de Entrenamiento con retropropagación

Paso 0: Iniciar los pesos y biases:

Establecer a valores pequeños y aleatorios

Paso 1: Mientras la condición de parada sea falsa, hacer pasos 2 – 9.

Paso 2: Para cada par entrada – salida de entrenamiento $s:t$, hacer pasos 3 - 8.

// FEEDFORWARD

Paso 3: Establecer la activación de cada unidad de entrada $i = 1, \dots, n$

$$x_i = s_i$$

y transmitiendo esta señal a todas las unidades de la capa oculta.

Paso 4: Calcular la entrada neta de cada unidad oculta $j = 1, \dots, p$

$$u_j = b_j + \sum w_{ji} \cdot x_i$$

Aplicar su función de activación para calcular su señal de salida

$$o_j = f(u_j)$$

Y enviar esta señal a todas las unidades de la capa siguiente.

Paso 5: Calcular la entrada neta de cada unidad de salida $k = 1, \dots, m$

$$u_k = b_k + \sum w_{kj} \cdot o_j$$

Aplicar su función de activación para calcular su señal de salida

$$y_k = f(u_k)$$

// RETROPROPAGACIÓN DE ERROR (BACKPROPAGATION)

Paso 6: Cada unidad de salida $k = 1, \dots, m$ recibe el valor deseado asociado al patrón de entrenamiento, y calcula su gradiente local.

$$\delta_k = (t_k - y_k) \cdot f'(u_k)$$

Calcula el término para la corrección de los pesos.

$$\delta w_{kj} = \eta \cdot \delta_k \cdot o_j$$

Calcula el término para la corrección de los biases.

$$\Delta b_k = \eta \cdot \delta_k$$

Y envía δ_k a las unidades de la capa anterior.

Paso 7: Cada unidad oculta $j = 1, \dots, p$ suma sus entradas delta (desde las unidades de la capa de salida)

$$\delta in_j = \sum \delta_k \cdot w_{kj}$$

Calcula el término de gradiente local

$$\delta_j = \delta in_j \cdot f'(u_j)$$

Calcula el término de corrección del peso

$$\Delta w_{ji} = \eta \cdot \delta_j \cdot x_i$$

Y el término de corrección de bias

$$\Delta b_j = \eta \cdot \delta_j$$

// ACTUALIZAR PESOS Y BIASES

Paso 8: Cada unidad de salida $k = 1, \dots, m$ actualiza su bias y sus

Pesos ($j = 0, \dots, p$)

$$w_{kj} (\text{nuevo}) = w_{kj} (\text{anterior}) + \Delta w_{kj}$$

$$b_k (\text{nuevo}) = b_k (\text{anterior}) + \Delta b_k$$

Cada unidad oculta $j = 1, \dots, p$ actualiza su bias y pesos ($i = 0, \dots, n$)

$$w_{ji} (\text{nuevo}) = w_{ji} (\text{anterior}) + \Delta w_{ji}$$

$$b_j (\text{nuevo}) = b_j (\text{anterior}) + \Delta b_j$$

Paso 9: Verificar la condición de parada

Si se logró un error menor al error cuadrático o se llegó al número de Epocas de entrenamiento entonces parar, sino continuar con paso 2.

Algoritmo de recuerdo

En esta etapa, después de haber encontrado los pesos ideales en el entrenamiento, se aplica sólo la fase Feedforward del mismo.

Paso 0: Establecer los pesos ideales (aplicar el algoritmo de entrenamiento)

// FEEDFORWARD

Paso 1: Para cada vector de entrada, hacer pasos 2 – 4

Paso 2: Para $i = 1$ hasta n , establecer la activación de la unidad de entrada x_i .

Paso 3: Para $j = 1$ hasta p

$$u_j = b_j + \sum w_{ji} \cdot x_i$$

$$o_j = f(u_j)$$

Paso 4: para $k = 1$ hasta m

$$u_k = b_k + \sum w_{kj} \cdot o_j$$

$$y_k = f(u_k)$$

**Constancia de participación de la tesis de maestría en Congreso Peruano de
Química**



**XXVIII Congreso Peruano de Química
"Dr. Gastón Pons Muzzo"
FQIQ – UNMSM
Marzo 29-31, 2017**



CONSTANCIA

La Dirección Científica deja constancia que el trabajo:

**APLICACIÓN DE REDES NEURONALES ARTIFICIALES (RNA) EN LA SORCIÓN
DE COBRE UTILIZANDO QUITOSANO ENTRECruzADO CON
GLUTARALDEHÍDO.**

De los autores:

Manuel Blanco Lecca, Christian Jacinto Hernández

Ha sido presentada al 28º Congreso Peruano de Química, realizado del 29 al 31 de
Marzo del año 2017 en la ciudad de Lima, Perú.

Lima, 31 de Marzo del 2017

Dra. Ana María Muñoz J.
Presidenta



Dra. Ana/C. Valderrama Negrón
Directora Científica

Sociedad Química del Perú
Av. Nicolás de Araníbar 696 Santa Beatriz- Lima 01
Teléfono: (511) 4723925 e mail: sqperu@gmail.com
www.sqperu.org.pe