

Construcción de una red neuronal para la simulación de la remoción de cromo hexavalente empleando carbón activado como adsorbente

Jorge Pellegrini¹, Jorge de Celis² y Juan Apesteguy³

¹ Universidad Tecnológica Nacional-Facultad Regional Avellaneda, Argentina,
jorgepellegrini@yahoo.com.ar

² Universidad Tecnológica Nacional-Facultad Regional Avellaneda, Argentina

³ Facultad de Ingeniería de la Universidad de Buenos Aires, Ciudad Autónoma de Buenos Aires, Argentina

RESUMEN

Se presenta el desarrollo de un modelo de red neuronal artificial como solución alternativa a la simulación de un proceso de adsorción. Para los ensayos de laboratorio se eligió el cromo hexavalente como contaminante y el carbón activado como material adsorbente.

Se evalúa el potencial que tiene la red neuronal artificial en comparación a los modelos matemáticos existentes en la actualidad.

Los resultados obtenidos indican que la red neuronal creada fue satisfactoria y explica con más exactitud la variabilidad del proceso.

Para la elección del entorno computacional y bibliotecas a utilizar se tuvo en cuenta por un lado que el software sea gratuito como también usar herramientas computacionales vanguardistas para que su aprendizaje sea enriquecedor para futuras investigaciones.

Palabras Clave: Adsorción, Red neuronal artificial, Tensorflow

INTRODUCCIÓN

Los procesos de adsorción presentan una importancia en el campo de la ingeniería ambiental en el tratamiento de contaminantes en solución acuosa. Las propiedades que presenta el carbón activado tales como una gran área superficial, micro porosidad, química superficial y bajo costo lo hace de uso recurrente para este campo de aplicación. En el presente trabajo, la síntesis del adsorbente es a partir de residuos ligno-celulósicos (cáscara de nuez) mediante activación química. Se eligió esta vía ya que se está aprovechando un residuo que de otra forma es un desecho de la industria agropecuaria.

El proceso de adsorción

Se trata de la remoción de cromo hexavalente de una solución líquida empleando un adsorbente carbonoso. El cromo se adsorbe en los sitios activos de la superficie del carbón activado. La cinética consiste en estudiar la evolución de la cantidad de contaminante adsorbido en función del tiempo. Expresar matemáticamente esa relación es importante con el fin de entender y predecir cómo el proceso evoluciona. Principalmente hay dos maneras de expresar la cinética:

-Modelo de primer orden

-Modelo de segundo orden

Los modelos consideran que la velocidad de cambio de la pendiente de la curva cinética cambia de forma lineal o cuadrática.

La adsorción en un medio poroso es un proceso por etapas. En el cual las resistencias predominantes del proceso van modificando su intensidad a medida que la superficie se va saturando.

El problema de los modelos vigentes es que consideran matemáticamente una resistencia media invariante de magnitud. Al no correlacionarse con la realidad empírica, que la resistencia predominante cambia con el tiempo, el ajuste a los datos no es del todo exacto.

A partir de esta falta se propone el empleo de una red neuronal que prediga la adsorción en función del tiempo, optimizándola para mejorar el desempeño frente a los modelos actuales.

Las redes neuronales son representaciones matemáticas que buscan replicar la mente humana teorizada y la actividad cerebral. Explorando la interconexión de un grupo de neuronas, se obtiene como resultado un poder de predicción de alto rendimiento de no linealidades entre las variables de un proceso.

Coulibaly et al. (2000) utilizan las redes neuronales para la predicción a largo plazo de las posibles entradas de energía para la planificación de operaciones hidroeléctricas. El autor concluye: " los modelos basados en redes neuronales proporcionan pronósticos más precisos que los modelos estocásticos tradicionales".

En forma simple, un sistema de redes neuronales artificiales consiste en capas de neuronas de entrada, ocultas y de salida. La capa de entrada recibe la información de las variables independientes, las capas ocultas aprenden la información y la capa de salida informa el valor de la variable dependiente. La forma del sistema se puede ver en la imagen 1.

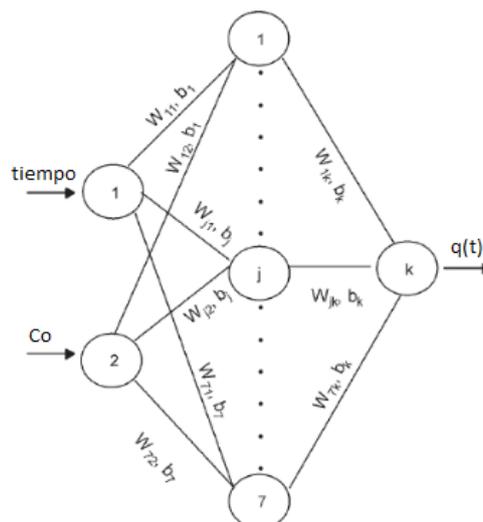


Figura 1. Arquitectura de la red neuronal.

Si la creación del sistema neuronal es adecuada, tiene la capacidad de aprender la información de la cinética y contará con la capacidad de ajustar a los datos. El rol fundamental lo tienen las capas ocultas ya que deben transferir la información de entrada a la salida, lo hace ajustando los pesos w_i de las neuronas (llamados pesos de conexión w , que representa la fuerza de conexión de las neuronas).

Objetivos

- Síntesis del material adsorbente a partir de un residuo lignocelulósico.
- Realización de ensayos de adsorción de cromo hexavalente en carbón activado.
- Ajuste matemático a los modelos cinéticos.
- Creación de una red neuronal artificial para el reconocimiento de la no linealidad de la cinética de adsorción. Que muestre una mejora considerable respecto a los otros modelos, siendo la creación de la misma mediante un software gratuito y simple de usar.
- Comparar los modelos utilizados.

METODOLOGÍA

Síntesis del adsorbente

Se colocó 10gr de cáscara en 20gr de ácido ortofosfórico 50% m/m. El sistema fue mezclado y posteriormente calentado durante 2h a una temperatura de 110°C y 1h a 400°C. Luego, el carbón fue lavado sucesivamente con agua destilada hasta pH constante.

Ensayo experimental

Para los ensayos de adsorción se pone en contacto a la solución de cromo con el carbón activado en erlenmeyers. La agitación se realiza en agitadores orbitales. Luego del tiempo del ensayo, la solución es filtrada y el cromo es medido por espectroscopia UV-visible empleando difenil-carbazida como

agente colorante. El tiempo de agitación será una variable para la obtención de información. Se define la capacidad de adsorción como la cantidad removida del contaminante por gramos del adsorbente.

$$q = \frac{(C_0 - C) * V}{m}$$

Donde q es la capacidad de adsorción [mg g^{-1}], C_0 y C son las concentraciones de soluto inicial y final [mg L^{-1}] del ensayo respectivamente y m/V es la dosis usada [mg L^{-1}].

La concentración de Cr (VI) utilizada fue de 50 mg/L con un dosaje de 0,1 gr/100ml. Para el estudio de la cinética se analizó el ajuste a los modelos de pseudo-primer orden y pseudo-segundo orden:

$$\frac{dq_t}{dt} = K_1 (q_e - q_t)$$

$$\frac{dq_t}{dt} = K_2 (q_e - q_t)^2$$

q_e y q_t son la capacidad de adsorción en el equilibrio y en el tiempo [mg g^{-1}], K_1 es la constante de adsorción de primer orden [min^{-1}] y K_2 es la constante de adsorción de segundo orden [$\text{g mg}^{-1} \text{min}^{-1}$]. Se define la tasa de adsorción como el cociente entre la capacidad de adsorción a un tiempo t y en el equilibrio:

$$F = \frac{q_t}{q_e}$$

Para poder comparar los resultados de forma consistente, se calcula la desviación estándar normalizada Δq :

$$\Delta q = \sqrt{\frac{\sum \left[\frac{[q_{t,exp} - q_{t,cal}]^2}{q_{t,exp}} \right]}{(n-1)}}$$

Creación de la red neuronal

OCTAVAS JORNADAS DE INGENIERÍA QUÍMICA SUSTENTABLE

Villa Dominico – 18, 19 y 20 de Septiembre 2018

Se denomina entrenamiento al proceso de aprendizaje de una red neuronal. Se opta por usar el método del descenso del gradiente ya que es el método más simple y también el más extendido y conocido. Consta de minimizar la pérdida (la diferencia cuadrática entre el valor estimado y el dato). Es un método de primer orden al hacer uso de un vector gradiente.

Python es el lenguaje programación elegido, ya que posee una sintaxis que favorece que el código sea legible. Por eso se procedió a aprender los elementos del lenguaje (bucles, funciones, clases, etc).

TensorFlow es una biblioteca de código abierto la cual se usará para la creación de la red neuronal. Desarrollada por Google, actualmente se utiliza tanto para la investigación dentro de la empresa como para la investigación mundial en aprendizaje automático ya que en noviembre de 2015 fue publicado bajo la licencia de código abierto. Que sea una tecnología de vanguardia utilizada en la actualidad fue uno de los principales motivos para su selección.

RESULTADOS

A continuación, se presenta el cuadro comparativo de la información obtenida para los modelos y la rutina del programa creado.

			Modelos				
			Red neuronal			1º orden	2º orden
qt [mg gr-1]	tiempo [min]	F	F -500 entrenamientos	F -1000 entrenamientos	F -10000 entrenamientos	F	F
8,25	2	0,622	0,710	0,692	0,635	0,001	0,064
8,38	4	0,632	0,710	0,693	0,638	0,002	0,121
11,03	128	0,832	0,742	0,745	0,792	0,049	0,816
11,40	246	0,860	0,771	0,790	0,874	0,086	0,896
12,33	1272	0,930	0,931	0,965	0,975	0,244	0,980
13,18	2712	0,995	0,982	0,992	0,983	0,285	0,992
13,26	4188	1,000	0,990	0,995	0,986	0,291	0,996

Perdida	0,030	0,022	0,004	3,472	0,577
Δq	0,0825	0,0982	0,0311	0,9348	0,4939
			R2	0,994	0,9984
			K1	0,000641	
			K2	0,000376	

Tabla 1. Resultados de los ajustes matemáticos a los modelos y desempeño de la red neuronal.

OCTAVAS JORNADAS DE INGENIERÍA QUÍMICA SUSTENTABLE

Villa Dominico – 18, 19 y 20 de Septiembre 2018

```
### importa tensorflow
import tensorflow as tf

### entrada de datos de la cinética a T=18°C
x= [0.00047752],[0.00095503],[0.03056108],[0.05873458],[0.30370076],[0.64751293],[1]]
y_ = [[0.62239873],[0.63185768],[0.83175103],[0.85985328],[0.92979904],[0.99450569],[1]]

### Creación de las capas
# 1x1 capa de entrada -> 1x2 capa oculta sigmoidal-> nx1 capa de salida con n= tamaño del vector x
# Capa 0 = entrada de datos x0, y0
x0 = tf.constant( x , dtype=tf.float32 )
y0 = tf.constant( y_ , dtype=tf.float32 )
# Capa 1 = 1x2 sigmoidal oculta
m1 = tf.Variable( tf.random_uniform( [1,2] , minval=0.1 , maxval=0.9 , dtype=tf.float32 ) )
b1 = tf.Variable( tf.random_uniform( [2] , minval=0.1 , maxval=0.9 , dtype=tf.float32 ) )
h1 = tf.sigmoid( tf.matmul( x0,m1 ) + b1 )
# Capa 2= nx1 salida
m2 = tf.Variable( tf.random_uniform( [2,1] , minval=0.1 , maxval=0.9 , dtype=tf.float32 ) )
b2 = tf.Variable( tf.random_uniform( [1] , minval=0.1 , maxval=0.9 , dtype=tf.float32 ) )
y_out = tf.sigmoid( tf.matmul( h1,m2 ) + b2 )
# Pérdida, como la sumatoria de la diferencia cuadrática entre el valor estimado y el dato
loss = tf.reduce_sum( tf.square( y0 - y_out ) )
# El método del gradiente descendente para el entrenamiento de la red
train = tf.train.GradientDescentOptimizer(1.0).minimize(loss)

# Entrenamiento
# 10 000 veces

# mostrar la estimación y coeficientes de la red
with tf.Session() as sess:
    sess.run( tf.global_variables_initializer() )
    for step in range(10000) :
        sess.run(train)
    results= sess.run([m1,b1,m2,b2,y_out,loss])
    labels = "m1,b1,m2,b2,y_out,loss".split(",")
    for label,result in zip(*(labels,results)) :
        print(int)
        print (label)
        print (result)
        print(int)
```

Figura 2. Rutina del programa

ANÁLISIS

Aunque los modelos de primer y según orden presentan un coeficiente de correlación R^2 alto, este no puede ser usado para comparación ya que no es un estadístico. Comparando la desviación estándar normalizada, el modelo

OCTAVAS JORNADAS DE INGENIERÍA QUÍMICA SUSTENTABLE

Villa Dominico – 18, 19 y 20 de Septiembre 2018

de primer orden explica el 7 % de la variabilidad de los datos, mientras que el modelo de segundo orden lo hace en un 51 % y la red neuronal en un 97%.

El modelo de primer orden no puede ajustarse a los datos por la propia naturaleza del proceso. Al ser el modelo lineal, la derivada de la curva del modelo no puede cambiar lo suficientemente rápido para ajustar a los datos, por eso el ajuste es tan pobre.

El modelo de segundo orden al ser cuadrático presenta una mejora en ese aspecto, obteniendo que los puntos finales llegan a asemejarse a los datos. El problema radica que en el principio del proceso hay un incremento muy alto de la adsorción en un tiempo muy corto debido al transporte externo. Ocasionando que el modelo no pueda ajustar a esos primeros puntos.

La red neuronal no presenta esos inconvenientes, teniendo un grado de explicación de la variabilidad mucho más alto y una pérdida de 100 ordenes menor respecto a los otros modelos.

Además, se estudió la cantidad necesaria de entrenamientos que necesita la red para alcanzar sus valores óptimos, siendo de 10 000 entrenamientos.

CONCLUSIONES

Se logró preparar carbón activado a partir de la cascara de nuez, siendo una materia prima renovable.

La red artificial desarrollada tuvo mejor desempeño que los modelos de primer y segundo orden. Esto es de importancia en la simulación de un lecho fijo ya que los primeros puntos de la cinética son los más importantes al estudiar el comportamiento de la torre (al funcionar mayoritariamente en estado transitorio).

Al aprender y manejar la arquitectura para crear redes neuronales con TensorFlow el futuro de la investigación es desarrollar un modelo de red neuronal para la múltiple funcionalidad y con esa información redefinir el gradiente acumulativo en las operaciones de adsorción en continuo.

REFERENCIAS BIBLIOGRÁFICAS

Allen, S.J.; McKay, G.; Khader, KYH. (1989). "Intraparticle diffusion of a basic dye during adsorption onto sphagnum peat", *Environ Pollut*, 56:39-50.

Coulibaly, P., F. Anctil, P. F. Rasmussen, and B. Bobfie, "A recurrent neural networks approach using indices of low-frequency climatic variability to forecast regional annual runoff", *Hydrol. Processes*, 14, 2755-2777, 2000.

Haykin, S. (2005). *Neural Networks*, 1. Editorial Macmillan College Publishing Company. EEUU. p.278.

NARENDRA K. S. and PARTHASARATHY K. "Identification and control of dynamical systems using neural networks". *IEEE Transactions on Neural Networks*, 1, 4-27, 1990.

Schmidt, S. y Launsby, R. (1994). "Understanding Industrial Designed Experiments". Editorial Air Academy Press. EEUU. pp.8- 23.

Wasserman, P. (1993). "Advanced Methods in Neural Computing". Editorial Van Nostrand Reinhold. EEUU. p.147.

Wu, F.C.; Tseng, R.L.; Juang, R.S. (2001). "Adsorption of Dyes and Phenols from Water on the Activated Carbons Prepared from Corn Cob Wastes", *Environmental Technology*, 22:2, 205-213.