

SOFTWARE

Sistema experto para el tratamiento de aguas residuales (SECTRARES)

Expert system for wastewater treatment (SECTRARES)

Lic. Yoan Martínez-López^I, Dr.C. María M. García^{II}, Dr.C. Rafael Bello Pérez^{II}, Dr.C. Rafael Falcón Martínez^{III}, Dr.C. Xiomara Cabrera Bermúdez^{II}

^IUniversidad de Camagüey, Facultad de Informática, Camagüey, Cuba.

^{II} Universidad Central de Las Villas, Santa Clara, Villa Clara, Cuba.

^{III} University of Ottawa, Electrical Engineering & Computer Science (EECS), Ottawa ON Canada.

RESUMEN. El tratamiento de las aguas residuales de una fábrica es de vital importancia, ya que elimina parte de la contaminación de las mismas, y contribuye con la protección del medio ambiente. Por lo cual resulta necesario llevar la experiencia adquirida por los especialistas del Instituto Virtual del Medio Ambiente de estos tratamientos, a distintas empresas del territorio, provincia y el país. El presente artículo consiste en la construcción de un Sistema Experto para el tratamiento de aguas residuales de una fábrica. La herramienta utiliza las Redes Neuronales Artificiales (RNA) para predecir el tipo de tratamiento que se debe aplicar.

Palabras clave: aguas residuales, redes neuronales, tratamiento.

ABSTRACT. The treatment of wastewater from a factory is of vital importance, since it eliminates some of the contamination of the same ones, and thus contributes to the protection of the environment. It is necessary to bring to different companies in the territory, province and country, the experience gained in these treatments by the specialists of the Virtual Institute of Environment. The present paper consists in building an Expert System for treating wastewater from a factory. The tool uses Artificial Neural Networks (ANN) to predict the type of treatment to be applied.

Keywords: wastewater, neural networks, treatment.

INTRODUCCIÓN

Las instalaciones industriales y de servicios generan numerosos residuales, los cuales terminan finalmente afectando al medio ambiente. Para evitar esta afectación ellos requieren un tratamiento dependiendo del tipo de desechos. Los residuos derramados sobre ríos, lagos y mares tienen propiedades contaminantes que se pueden transmitir al receptor final. De modo que es necesario estudiar sus características físicas, químicas y biológicas.

En este contexto el Instituto Virtual de Medio Ambiente de la UCLV (Universidad Central Marta Abreu de Las Villas) planteó la importancia de extender la experiencia adquirida por los especialistas de su área a distintas empresas del territorio, provincia y país. Se trata de poder pronosticar sin la presencia de un experto que tratamiento debían recibir las aguas residuales vertidas por distintas fábricas, conociendo un grupo de parámetros y explicando el porqué de la solución que se alcance en cada momento

(Rodríguez, 1996; Rodríguez, 2007).

Sistemas similares fueron desarrollados en nuestro país para el tratamiento de aguas residuales, como es el caso del modelo híbrido aplicado para el tratamiento de aguas residuales (Rodríguez, 2005; Cabrera *et al.*, 2009), entre otros modelos (Herrera, 1996). Asimismo existen varios sistemas informáticos que fueron desarrollados en el resto del mundo, como son los sistemas informáticos para el diseño de Estaciones de Tratamiento de Aguas Residuales (ETAR), que resultan de gran utilidad para la toma de decisiones relacionadas con la gestión de la calidad del agua en el medio natural, como es el caso del Diseño Automatizado de Tratamientos de Aguas Residuales (DATAR), desarrollado para el diseño completo y riguroso de una ETAR que cumpla las especificaciones impuestas al vertido (Ferrer *et al.*, 1994) y el sistema cad/cae para plantas de tratamiento de aguas residuales (Gasso, 1989).

Se contaba con un conjunto de ejemplos correspondiente a tratamientos desarrollados por los expertos, caracterizados cada ejemplo por descripciones de los parámetros de las aguas en vertimiento. Los rasgos eran de dominios diferentes y existían rasgos de valores múltiples.

El problema se reduce a determinar la secuencia óptima de tratamientos que debe aplicarse a las aguas residuales vertidas sobre una fuente de agua no contaminada.

A partir de la información caracterizada por estos parámetros se conforma una Base de Conocimientos correspondiente a tratamientos con respecto a los parámetros que caracterizaban las aguas residuales. Como propuesta de solución se propone una red neuronal con una topología de una sola capa.

MÉTODOS

Construcción de la Base de Conocimientos

Para la resolución del problema del tratamiento de aguas residuales, los expertos consideran las características de las aguas residuales que se enumeran en la Tabla 1; todas ellas son consideradas como atributos predictores del problema y poseen valores continuos.

El último atributo se refiere al tratamiento a llevar a cabo, lo cual se especifica a través de una secuencia de etapas del proceso. Dicha especificación es una tarea que también requiere del conocimiento experto. Los expertos humanos tienen que seleccionar las etapas apropiadas y el orden de aplicación de las mismas. En consecuencia, para cada problema en este atributo se especifica la secuencia de etapas que conforman el tratamiento adecuado, aunque no el orden de aplicación de estas. Esto convierte al atributo “tratamiento” en un atributo multi-evaluado. En total son 11 atributos predictores y un atributo objetivo.

La solución parcial brindada por la red se complementa con una componente basada en casos que recupera los casos más similares (soluciones previas) al problema actual. De esta forma, el usuario puede estimar, en el contexto de las soluciones ofrecidas a los casos recuperados, cuál debe ser el orden de las etapas propuestas por las Redes Neuronales Artificiales (RNA) (Freeman & Skapura, 1993).

En este enfoque híbrido, la RNA permitió extraer el conocimiento, hacer las inferencias de los tratamientos, así como explicar el porqué de los resultados obtenidos al usuario; complementando la capacidad del sistema inteligente en la toma de decisiones.

TABLA 1. Atributos de la Base de Conocimiento para el tratamiento de las aguas residuales

Atributo	Interpretación	Valores
BOD5	Demanda biológica de oxígeno	Baja, Ligeramente baja, Media, Alta
COD	Demanda química de oxígeno	Baja, Alta
BOD/COD	Razón BOD/COD	Baja, Media, Alta
pH	Acidez del agua al abandonar la planta	Ácida, Neutra, Base
N	Compuestos nitrogenados	Bajo, Alto
P	Compuestos fosfóricos	Bajo, Alto
ST	Los residuos luego de la evaporación	Bajo, Alto
STV	Sólidos totales volátiles	Bajo, Alto
SDT	Sólidos filtrables	Bajo, Alto
SDV	Sólidos volátiles disueltos	Bajo, Alto
SS	Sólidos suspendidos	Bajo, Alto

Análisis de los modelos de RNA a emplear

La posibilidad de inferir en otro momento distintos rasgos objetivos sugirió emplear una topología de una sola capa. Ante esta situación se valoró el modelo Hopfield, (Hopfield, 1982), o el modelo de red neuronal de activación interactiva y competencia (IAC), (McClelland, 1989), reportándose mejores resultados con este último o su variante SIAC.

El diseño de la topología de la RNA es de la siguiente manera: Se crea un grupo de neuronas por cada atributo de la Base de Casos y se incorporan en el grupo tantas neuronas como valores posibles existen del rasgo. El aprendizaje en este modelo de RNA es no adaptativo, los pesos se calculan empleando el coeficiente de correlación de Pearson (Rodgers & Nicewander, 1988), aunque pudieran usarse otras variantes como frecuencia relativa.

Para la inicialización de los vectores de entrada externa se calcula el grado de pertenencia del valor dado al rasgo con relación a los distintos términos lingüísticos definidos. Se activará

la neurona que representa al término lingüístico de “máxima pertenencia” (Zhou *et al.*, 1997). En la Figura 1 se muestra la topología de la RNA a emplear.

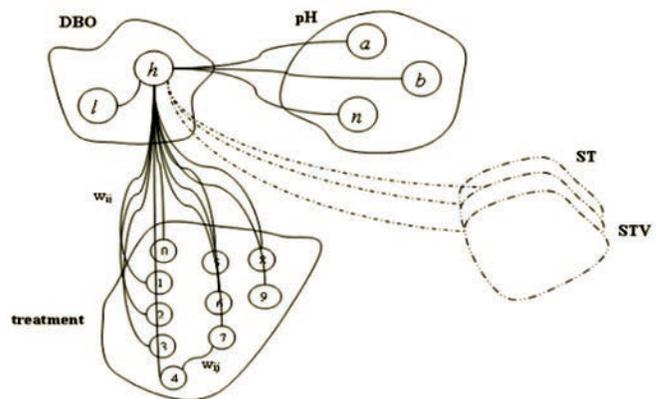


FIGURA 1. Topología de la RNA empleada para el tratamiento de aguas residuales.

Propuesta del módulo razonador basado en casos

El módulo razonador consiste en la búsqueda de causas de un hecho, es decir según la opinión científica es la deducción y comprensión de leyes y condiciones que describan un proceso; para los Sistemas Expertos, es dotar a la herramienta de explicación de porqué el sistema ha obtenido un resultado dado o como ha llegado a dicha conclusión, con el fin de ayudar al especialista a comprender las respuestas del mismo.

La organización de la Base de Casos y la función de disimilitud

La recuperación se realiza utilizando una medida de similitud que compara el problema resuelto (P) con cada caso (R) de la Base de Casos, tomando valores en el intervalo [0,1]. Teniendo en cuenta este criterio y la siguiente expresión según (García, 1997; Martínez, 2010):

$$D(P, R) = \sum_{C \in C_p} d(P.c, R.c) \quad (1)$$

$$d(P.c, R.c) = \begin{cases} 0 & \text{si } P.c = R.c \\ W_c * S_c + 0.1 & \text{en caso contrario} \end{cases} \quad (2)$$

$$W_c = \begin{cases} 0 & \text{si } ORDEN(P.c, C) = 0 \\ \sum_{j=1}^{NED(K)} W_{cj} & \text{en caso contrario} \end{cases} \quad (3)$$

$$W_{cj} = \begin{cases} MRR[c, K].P[P.c, j] & \text{si } ORDEN(P.K, K) \neq j \\ (MRR[c, K].P[P.c, j])^2 & \text{si } ORDEN(P.K, K) = j \end{cases} \quad (4)$$

$$S_c = \sqrt{\sum_{j \in C_o} (MRR[c, j].P[P.c, P.j] - MRR[c, j].P[R.c, P.j])^2} \quad (5)$$

donde:

P-denota el patrón de completado;

R-denota el registro de la Base de Casos al cual se le calcula su desigualdad con el patrón;

Cp-denota el conjunto de rasgos predictores;

P.c y R.c-denotan el valor dado al rasgo C en el patrón P y el registro R respectivamente;

K-denota el rasgo objetivo cuyo valor fue inferido con máxima activación.

MRR-matriz de relación entre rasgo;

ORDEN (e,i)-la cual retorna el orden del valor e para el rasgo i;

NED (K)-denota la cantidad de elementos del dominio del rasgo K;

W_c y S_c -importancia y equivalencia de los rasgos predictores para inferir el rasgo objetivo;

Co-denota el conjunto de rasgos objetivos.

RESULTADOS Y DISCUSIÓN

Construcción del prototipo del sistema

Tomando en consideración todo lo anteriormente explicado, se procedió a construir una aplicación final que reutilizó los componentes ofrecidos por la plataforma Neuro Evaluator (Falcón, 2003; Falcón, 2006).

A partir de la plataforma NeuroEvaluator se obtuvo un prototipo que se independizó de la plataforma, derivándose el software "Sistema Experto Conexionista para el Tratamiento de Aguas Residuales"

En el software, el usuario puede: Conformar su propia Base de Casos mediante una ventana de captura de datos (Figura 1), o emplear alguna ya existente.

Se definen las muestras de aprendizaje y control como paso previo para entrenar la RNA y posteriormente validar el nivel de generalización alcanzado por la misma. Se le da la posibilidad al usuario de dar un solo fichero y en ese caso

seleccionar las muestras de aprendizaje y control a partir de un umbral con el 70% de los datos como muestra de aprendizaje, sobre 10 iteraciones).

Es válido destacar que la implementación de la aplicación final comprendió diseñar la interfaz de usuario e invocar las siguientes bibliotecas de enlace dinámico provenientes de NeuroEvaluator:

SIACNetwork.dll, se implementó el modelo de SIAC: Al ser el modelo de RNA que muestra mayor nivel de precisión y efectividad, dado a las características que presenta este modelo.

Preprocessor.dll, que encapsula el módulo de pre-procesamiento de las RNA.

También fue necesario implementar la componente basada en casos (RBC) para explicar la solución ofrecida por la RNA.

Requerimientos del software

El software requiere para su funcionamiento, tener instalado el sistema operativo Windows 2000, o versiones superiores

y el fichero "Aguas_Residuales.exe", además de la Base de Casos conformada con los ficheros (*.data, *.prep, *.pool), donde:
 *.data, es el fichero almacena la Base de Casos.
 *.prep, es el fichero de pre-procesamiento de las RNA.
 *.pool, es el fichero de almacenamiento de las neuronas de las RNA.

La herramienta consta de un menú principal referidos a: **Opciones, Inferir tratamiento e Información del software.**
 El menú **Opciones**, permite realizar un conjunto de acciones sobre la **Base de Casos** tales como: **Cargar una Base de Casos** ya creada, **Crear Base de Casos** o **Cerrar la aplicación:**

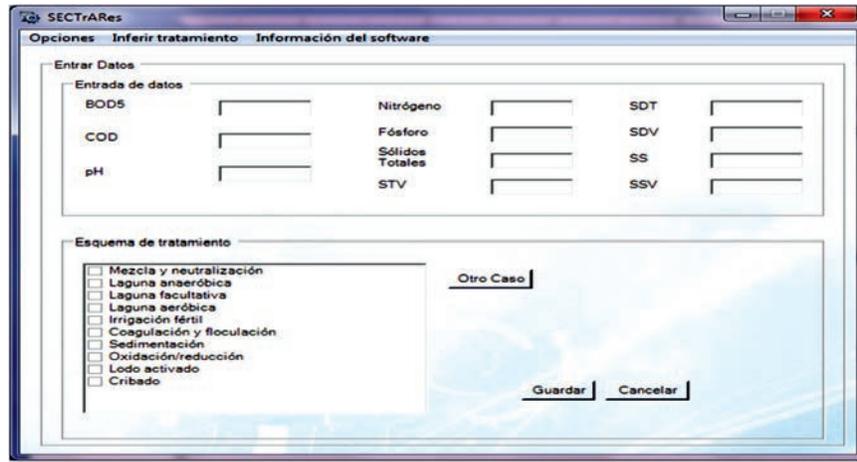


FIGURA 2. Ventana para conformar la Base de Casos.

Para Inferir tratamiento, el usuario deberá teclear los rasgos predictores relativo a las aguas para poder inferir el tratamiento a seguir y obtendrá una explicación del porqué de la solución encontrada a partir de listar los K casos más similares.

Luego de entrar los datos a la aplicación (BD5, COD, etc.) y de seleccionar la cantidad de casos similares, el experto podrá obtener la Respuesta de la RNA para esta instancia Figura 2.

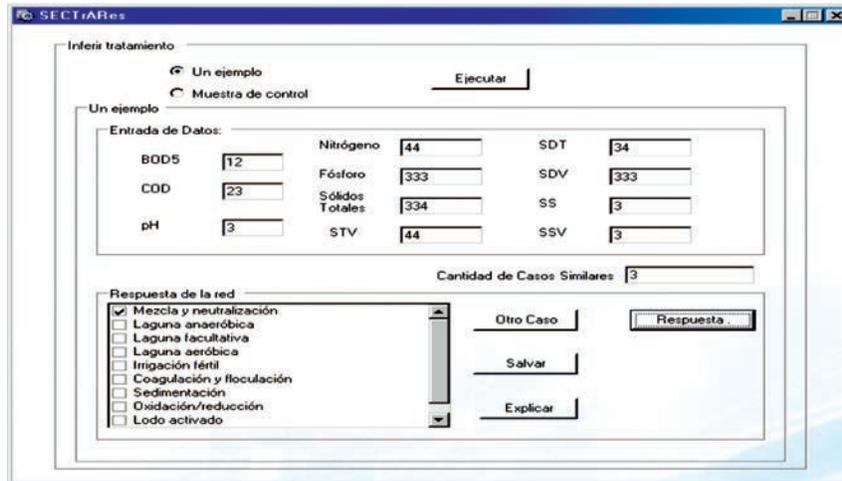


FIGURA 3. Inferencia de la RNA para el tratamiento de aguas residuales.

Por último si desea saber la explicación de los casos más similares, la herramienta permite esa facilidad, al presionar el botón Explicar, Figura 4.



FIGURA 4. Casos más similares.

También la herramienta le brinda al usuario la operación de guardar este nuevo caso, lo que permite el autoaprendizaje de la RNA, al presionar el botón Salvar.

AGRADECIMIENTOS

Al Dr.C. Yovani Marrero Ponce, Lic. en Farmacia, Profesor de Facultad de Química Farmacia, de la Universidad Central de Las Villas, y a todos los que han colaborado de una manera u otra en la construcción del Sistema Experto para el Tratamiento de Aguas Residuales.

REFERENCIAS BIBLIOGRÁFICAS

- CABRERA, X.; CONTRERAS, A.M.; HERRERA, Z.: *Influencia de la caracterización de los residuales líquidos industriales en la selección del sistema de tratamiento más adecuado*, Revista *CentroAzucar*, Ed. Universidad Central de las Villas (UCLV), Santa Clara, Villa Clara, Cuba, 2009.
- FALCÓN, M. R. J.: *Herramienta para el desarrollo de sistemas conexionistas con presencia de rasgos difusos*, Ed. Universidad Central de las Villas (UCLV), Santa Clara, Villa Clara, Cuba, 2003.
- FALCÓN, M. R. J.: *Evaluador de modelos de RNA asociativas con manejo de rasgos borrosos*, Tesis (en opción al título de Máster en Ciencia de la Computación), Universidad Central de las Villas (UCLV), Santa Clara, Villa Clara, Cuba, 2006.
- FERRER, J.; GABALDÓN, C.; SECO, G.; MARZAL, P.: "Utilización de sistemas informáticos para el diseño de estaciones de tratamiento de aguas residuales", [en línea] *Revistes i congressos UPC*, 1 (4), 1994. Disponible en: <http://upcommons.upc.edu/revistes/handle/2099/4208>[Consulta: mayo 18 2010].
- FREEMAN, J. A. & D. SKAPURA.: *Redes neuronales: algoritmos, aplicaciones y técnicas de programación*, Addison-Wesley Iberoamericana, New York, USA, 1993.
- GARCÍA, L. M. M.: *El empleo del razonamiento basado en casos en el desarrollo de sistemas basados en el conocimiento para el diagnóstico*, Ed. Universidad Central de las Villas (UCLV), Santa Clara, Villa Clara, Cuba, 1997.
- GASSO D. S.: *Desarrollo de un sistema CAD/CAE para plantas de tratamiento de aguas residuales*, Tesis (en opción al grado científico de Doctor en Ciencias Informáticas), Universidad Politécnica Catalunya, España, 1989.
- HERRERA, Z.: *Complex Processes Analysis Application to the problem of contamination for wastewater sugar industry*, Tesis (en opción al título de Máster en Análisis de Procesos en la Industria Química), Universidad Central de Las Villas, Cuba, 1996.
- HOPFIELD, J.J.: "Neural networks and physical systems with emergent collective computational abilities", *Proc. Nat. Acad. Sci*, 79: 2554–2558, 1982.
- MARTÍNEZ, L. Y.: *Shell para la construcción de Sistemas Expertos Conexionistas*, Trabajo de Diploma (en opción al título de Licenciado en Ciencias de la Computación), Universidad Central de las Villas (UCLV), Santa Clara, Villa Clara, Cuba, 2010.
- MATICH, D. J.: *Redes neuronales, Conceptos básicos y aplicaciones*. Ed. Universidad Tecnológica Nacional, México, marzo, 2001.
- MCCLELLAND, J. L. R.: *Explorations in parallel distributed processing*, MIT Press, Cambridge, USA, 1989.
- RODGERS, J. L. & NICEWANDER, W. A.: "Thirteen ways to look at the correlation coefficient", *The American Statistician*, 42 (1): 59–66, 1988.
- RODRÍGUEZ, S. Y.: *Generalización de la Métrica Basada en la Diferencia de Valores para Variables Lingüísticas y su Aplicación en Sistemas Basados en el Conocimiento*, Tesis (en opción al grado científico de Doctor en Ciencias de la Computación) Universidad Central de las Villas (UCLV), Santa Clara, Villa Clara, Cuba, 2007.
- RODRIGUEZ, S. Y.: *Sistema Basado en Casos para el diagnóstico de malformaciones cardiovasculares en recién nacidos*, Trabajo de Diploma (en opción al título de Licenciado en Ciencia de la Computación), Universidad Central de las Villas (UCLV), Santa Clara, Villa Clara, Cuba, 1996.
- RODRÍGUEZ S. Y.: CBR-ANN hybrid model to optimize the sequence of wastewater treatments. In: Second International ICSC Symposium on Information Technologies in Environmental Engineering, Shaker Verlag 2005, pp. 711-720, ISBN 3-8322-4362-3, Otto-von-Guericke-Universität Magdeburg, Germany, 2005.
- ZHOU, Q., PURVIS, M., & KASABOV, N.: Membership Function Selection Method for Fuzzy Neural Networks, Dunedin, New Zealand, Proceedings of the International Conference on Neural Information Processing and Intelligent System, New Zealand, 1997.

CONCLUSIONES

- Con la implementación de esta herramienta, el experto o cualquier usuario con un mínimo de conocimiento sobre el tratamiento de aguas residuales, podrá obtener uno o varios posibles tratamientos, dada las características químicas del agua que se esté analizando. Aunque en estos momentos la herramienta no está en uso, se considera que debe ser explotada, ya que un sistema desarrollado con este enfoque da respuesta al problema planteado anteriormente.

Recibido: 28 de febrero de 2013.

Aprobado: 9 de julio de 2014.

Yoan Martínez López, Lic. en Ciencias de la Computación, Profesor de la Universidad de Camagüey, Facultad de Informática, Circunvalación Norte km 51/2, Camagüey, Cuba, CP: 74650, Correo electrónico: yoan.martinez@reduc.edu.cu

Nota: La mención de marcas comerciales de equipos, instrumentos o materiales específicos obedece a propósitos de identificación, no existiendo ningún compromiso promocional con relación a los mismos, ni por los autores ni por el editor.