

Investigación actual del grupo SOCO: Metodologías híbridas de *Soft Computing*

René Alquézar, Lluís Belanche, Àngela Nebot, Enrique Romero y Alfredo Vellido

Grup de Soft Computing

<http://www.lsi.upc.es/dept/investigacion/sectia/soco/>

Secció d'Intel·ligència Artificial, Dept. Llenguatges i Sistemes Informàtics

Universitat Politècnica de Catalunya

Barcelona

{alquezar, belanche, angela, eromero, avellido}@lsi.upc.es

Abstract. El presente informe detalla la actividad de investigación del grupo de *Soft Computing (SOCO)*, adscrito al departamento de Lenguajes y Sistemas Informáticos (LSI) de la Universidad Politécnica de Cataluña (UPC). Se centra en una descripción general de las actuales líneas de investigación y una descripción detallada de los resultados más recientes, con especial énfasis en las técnicas y modelos desarrollados y su aplicación en diversas áreas del conocimiento.

1 Introducción

Soft computing es un término acuñado a mediados de la década de los 90 para describir el uso combinado de diferentes aproximaciones computacionales surgidas en los últimos treinta años y entre las que destacan los sistemas difusos, las redes neuronales y los algoritmos evolutivos. Adolece de cierta vaguedad y acaso se beneficiaría de cierta contextualización: *Soft Computing* comparte elementos, entre otros, de *Machine Learning (ML)*, *Artificial Intelligence (AI)*, *Data Mining (DM)*, *Knowledge Discovery in Databases (KDD)* y *Neural Computing (NC)*.

Pese a sus evidentes diferencias, el común denominador de estas metodologías es su abandono de la lógica binaria, los modelos analíticos estáticos, las clasificaciones rígidas y las búsquedas deterministas. Al abordar problemas del mundo real, es habitual el que estos estén poco y/o mal definidos, siendo difíciles de modelar. En estos casos, los modelos precisos son poco prácticos, demasiado costosos o simplemente inexistentes.

El grupo de *Soft Computing (SOCO)* forma parte de la sección de Inteligencia Artificial del departamento de Lenguajes y Sistemas Informáticos (LSI) de la Universidad Politécnica de Cataluña (UPC). Está compuesto por tres profesores titulares de universidad (TU), Àngela Nebot, René Alquézar y Lluís Belanche, un profesor asociado a tiempo completo, Enrique Romero, y un becario Ramón y Cajal, Alfredo Vellido. Actualmente se están dirigiendo 7 tesis doctorales.

El resto del informe se estructura como sigue: en primer lugar, se describen, de forma general, las principales líneas de investigación del grupo, asociadas a técnicas y modelos determinados. Tras ello, se da cuenta de los resultados más recientes obtenidos

en el desarrollo de estas líneas y la aplicación de los modelos y técnicas, producto de las mismas, en diferentes ámbitos. Posteriormente se comentan los avances en el desarrollo de hibridaciones de los métodos anteriormente propuestos. Por último, se apuntan líneas de trabajo de reciente apertura y se facilitan las bibliografías tanto general como producto directo de las actividades del grupo de investigación.

2 Líneas de investigación: descripción general

La actividad investigadora del grupo en el área de las técnicas y sistemas híbridos de *soft computing* comprende:

- **Selección de variables relevantes y reducción de la dimensión.** Al utilizar una gran cantidad de variables en un proceso de aprendizaje inductivo, existe la posibilidad de que algunas no aporten suficiente información útil para el proceso. Por ello es fundamental reducir el espacio de hipótesis en un intento de encontrar subconjuntos de variables que den lugar a un mejor rendimiento de los algoritmos de inducción. Esta línea está coordinada por el doctor Ll. Belanche.
- **Sistemas basados en técnicas difusas.** Se ha trabajado en estos sistemas desde la perspectiva del modelado y simulación de sistemas dinámicos, desarrollándose la metodología de *Razonamiento Inductivo Difuso (FIR)*. Esta línea está coordinada por la doctora A. Nebot en colaboración con el doctor F.E. Cellier de la *University of Arizona*.
- **Sistemas conexionistas.** En esta línea se ha centrado la investigación en las Redes Neuronales Recurrentes (*RNN*) y las *Feed-forward (FNN)*. Los temas más estudiados por el grupo han sido la relación entre aprendizaje y generalización, el uso de diferentes funciones de activación, el estudio de modificaciones en las funciones de coste, el pre/post-procesado de los datos, así como la formalización y adaptación de diferentes algoritmos de entrenamiento. Esta línea está coordinada por el doctor R. Alquézar y el profesor Enrique Romero. El grupo tiene también experiencia en sistemas conexionistas híbridos que combinan técnicas conexionistas, genéticas y difusas. Se han desarrollado las Redes Neuronales Heterogéneas (RNH) basadas en la generalización del modelo clásico de neurona. La línea está coordinada por el doctor Ll. Belanche en colaboración con el doctor J.J. Valdés del *National Research Council* de Canadá.
- **Sistemas probabilísticos.** Esta es una línea de reciente incorporación a los trabajos del grupo. Conciernen la definición de modelos habitualmente asociados a las áreas de inteligencia artificial y *machine learning* dentro de un marco formal probabilístico, desde *Feed-forward Neural Networks (FNN)* hasta modelos no supervisados como el *Generative Topographic Mapping (GTM)*. Esta línea es desarrollada por el doctor Alfredo Vellido y en ella se colabora con el catedrático P.J.G. Lisboa y con el doctor W. El-Deredy, de la *John Moores University* de Liverpool, U.K.

El trabajo desarrollado por el grupo desde sus inicios se enmarca en dos proyectos de investigación CICYT:

- Proyecto TAP99-0747, “Metodologías híbridas para el análisis de fallos en sistemas dinámicos complejos”, realizado por el Departamento de LSI de la UPC, conjuntamente con el departamento de ESII, el departamento de IOE y el Instituto de Robótica e Informática Industrial. *Duración: 1999-2002.*
- Proyecto DPI2002-03225, “Metodologías híbridas de soft computing y su aplicación a la detección de anomalías en estaciones depuradoras de aguas residuales”, realizado por el el Departamento de LSI de la UPC, conjuntamente con el departamento de ESII, el Instituto de Robótica e Informática Industrial y la Universidad de Girona, partiendo del trabajo desarrollado anteriormente. *Duración: 2003-2006.*

3 Líneas de investigación: metodologías de *Soft-Computing*

3.1 Selección de variables relevantes y reducción de la dimensión

Se ha realizado una exhaustiva revisión de los algoritmos de selección de variables en la literatura de *machine learning*, así como un estudio comparativo en problemas artificiales controlados. Para ello se ha desarrollado un *software* para selección de variables hasta ahora inexistente en la literatura, compuesto por un generador de datos artificiales, un simulador de algoritmos y un evaluador automático de la calidad de las soluciones [Molina *et al.*, 2002]. En base a las conclusiones obtenidas, se trabaja actualmente en un simulador de algoritmos secuenciales dotado de diversos criterios de evaluación, búsqueda y parada.

3.2 Razonamiento Inductivo Difuso (*Fuzzy Inductive Reasoning – FIR*)

El Razonamiento Inductivo Difuso (*Fuzzy Inductive Reasoning - FIR*), es una metodología cualitativa basada en modelos que se deriva de la Teoría General de Sistemas y que resulta atractiva para el modelado y predicción de sistemas dinámicos.

Algunas de las principales ventajas del *FIR* son:

- El razonamiento inductivo permite tratar el tiempo en los modelos cualitativos como una variable cuantitativa (continua).
- La técnica es de aplicación a cualquier sistema accesible a la experimentación y observación. El razonamiento inductivo se basa completamente en el comportamiento no siendo necesario conocimiento de la estructura interna del sistema.
- La metodología contiene un mecanismo inherente de validación en su motor de simulación cualitativa lo que impide llegar a conclusiones no justificables por los hechos acontecidos.

Se ha desarrollado esta metodología en colaboración con la *University of Arizona* [Nebot *et al.*, 1994; Múgica y Nebot, 1996; Nebot y Múgica, 1996; Escobet *et al.*, 2000] *FIR* ha servido como base para el desarrollo de un Sistema de Monitorización de Fallos mejorado [Escobet *et al.*, 1999a; 1999b].

3.3 Redes Neuronales Artificiales

3.3.1 Redes Neuronales *Feed-Forward*

Uno de los primeros temas estudiados por *SOCO* fue la utilización de neuronas con función de activación sinusoidal y sus efectos sobre la capacidad y velocidad de aprendizaje de las redes. Se ha experimentado con las capas *feed-forward* de una red recurrente aumentada (*ASLRNN*) [Sopena y Alquézar, 1994]. También se han estudiado las propiedades de la función de activación sinusoidal y se argumenta en favor de su uso para redes *feed-forward* [Alquézar, 1997; Sopena *et al.*, 1999].

Por otro lado, el grupo ha trabajado en la aplicación de redes *feed-forward* del tipo *time-delay* a problemas de predicción de señales biomédicas [Cueva *et al.*, 1997; Alquézar *et al.*, 1998]. También se ha estudiado el uso de técnicas markovianas de pre-procesado de datos para reducir las entradas y pesos de una red *time-delay* [Alquézar *et al.*, 1998].

Se ha propuesto un método incremental de aproximación de funciones denominado *SAOCIF* (*Sequential Approximation with Optimal Coefficients and Interacting Frequencies*), que puede implementarse mediante la construcción neurona a neurona de una red neuronal *feed-forward* [Romero, 2000; Romero y Alquézar, 2002a]. Asimismo, se ha propuesto una metodología de utilización del método de selección de variables *Sequential Backward Selection* con perceptrones multicapa [Romero *et al.*, 2003], con aplicaciones diversas en psicología [Armadians *et al.*, 2003] y medicina [Selva *et al.*, 2002, Selva *et al.* 2003, Sopena y Romero, 2004].

La utilización de métodos Bayesianos para la óptima regularización de redes *feed-forward* multicapa y para la selección de variables de entrada ha sido estudiada en [Vellido *et al.*, 2000a; Lisboa *et al.*, 2000; Vellido y Lisboa, 2001].

3.3.2 Redes Neuronales Recurrentes

La investigación realizada por el grupo en el área de las redes neuronales recurrentes se ha centrado en el estudio teórico de modelos de *RNNs* de tiempo discreto [Sanfeliu y Alquézar, 1992; Alquézar y Sanfeliu, 1993, 1995] y en su aplicación a problemas de inferencia gramatical (para reconocimiento de secuencias) [Alquézar y Sanfeliu, 1994a, 1994b; Sopena y Alquézar, 1994; Alquézar *et al.*, 1997] y predicción de sistemas [Cueva *et al.*, 1997]. El trabajo más relevante dentro de esta línea consistió en un análisis de las capacidades de las redes *SLRNNs* y *ASLRNNs* para representar máquinas de estados finitos [Alquézar y Sanfeliu, 1993, 1994b, 1995].

Por otro lado, se ha estudiado la influencia de diferentes funciones de activación en el rendimiento de las redes recurrentes, en [Alquézar y Sopena, 1994; Sopena y Alquézar, 1994; Alquézar y Sanfeliu, 1994a]. El grupo ha trabajado también en la adaptación del algoritmo de Schmidhuber [Schmidhuber, 1992] a redes *SLRNN* de segundo orden, y se formalizó el algoritmo de entrenamiento que calcula el gradiente real en redes *ASLRNN* [Alquézar, 1997]. En cuanto a la aplicación de las *RNNs* al modelado y predicción de sistemas, se han usado *ASLRNNs* para aprender modelos del sistema nervioso central de control [Cueva *et al.*, 1997; Alquézar *et al.*, 1998].

3.3.3 Redes Neuronales Heterogéneas

El modelo de neurona heterogénea fue propuesto en [Valdés y García, 1997] y posteriormente desarrollado por investigadores de SOCO [Belanche *et al.*, 1998; Belanche y Valdés, 1998; Belanche, 2000]. La neurona heterogénea es una generalización del modelo clásico de neurona que permite trabajar de manera más natural con información heterogénea, incompleta y/o imprecisa. Estas neuronas heterogéneas pueden trabajar directamente con números difusos y complejos, además de con variables discretas. Adicionalmente, se permite la presencia de valores perdidos en cualquiera de estos dominios. Las neuronas heterogéneas pueden combinarse en diferentes topologías y arquitecturas, dando lugar a redes neuronales heterogéneas (*Heterogeneous Neural Networks - HNN*). Se ha desarrollado una extensión de red de Kohonen que puede tratar variables heterogéneas y valores perdidos [Negri y Belanche, 2001]

Las redes heterogéneas requieren el desarrollo de algoritmos de aprendizaje diferentes de los ya establecidos para las redes clásicas. Se ha propuesto el uso de algoritmos genéticos *breeder (BGA)*, así como un método de entrenamiento basado en la estrategia de “divide y vencerás” [Quartier y Belanche, 2001; Belanche, 1999].

3.4 Modelos probabilísticos no supervisados

SOCO ha comenzado recientemente a investigar en el área de modelos no supervisados para el agrupamiento y visualización de datos de dimensión alta, con especial énfasis en aquellos formulados dentro de un marco probabilístico. El *Generative Topographic Mapping (GTM)*, [Bishop *et al.*, 1998]) es un modelo generativo no supervisado, no lineal, de variables latentes, que responde a tales características. Está inspirado en los bien conocidos *Self-Organizing Maps (SOM)*, pero a diferencia de estos su desarrollo es completamente probabilístico. El *GTM* mantiene las muchas ventajas del *SOM*, a un tiempo que resuelve muchas de sus limitaciones gracias a su definición en un marco probabilístico. *GTM* puede también ser usado en la reconstrucción estadística de datos incompletos (*missing data*) ofreciendo mejores resultados que otros métodos alternativos. Se están actualmente aplicando estas técnicas y otros desarrollos paralelos [Vellido *et al.*, 2003], en el proyecto STREAMES, al agrupamiento de cuencas fluviales y la regeneración de datos perdidos en un conjunto de datos de alta dimensionalidad, muy incompleto y con bajo número de observaciones [Vellido *et al.*, 2004; Vicente *et al.*, 2004].

4 Líneas de investigación: Hibridaciones de métodos y técnicas de *Soft-Computing*

Actualmente el grupo SOCO está investigando en la hibridación de las diferentes metodologías desarrolladas con el objetivo de mejorar su rendimiento y fiabilidad. En esta sección se apuntan las 6 sublíneas en las que se está trabajando:

4.1 Redes Neuronales *Feed-Forward* y *Support Vector Machines*

Se ha trabajado en SOCO en la simulación del comportamiento de una *SVM* con una *FNN*. El progreso en esta línea de investigación se ha plasmado en el diseño de un algoritmo de maximización del margen con *FNNs* no restringido a funciones *kernel* ni a frecuencias en el conjunto de aprendizaje [Romero y Alquézar, 2002b]. Este algoritmo se ha aplicado en problemas del área del procesamiento del lenguaje natural, comparando varios clasificadores basados en el margen (*SVMs*, *AdaBoost* y el modelo mixto *FNN-SVM* con ponderación del error cuadrático) [Romero *et al.*, 2004].

4.2 Razonamiento Inductivo Difuso, *Simulated Annealing* y Algoritmos Genéticos

Se han diseñado e implementado algoritmos basados en técnicas de *simulated annealing* y en algoritmos genéticos, en el contexto de *FIR*, para la determinación de particiones óptimas de una manera automática [Nebot, 2003; Acosta *et al.*, 2004]. Estos modelos han sido aplicados a la predicción de contaminación ambiental en la ciudad de México y a la predicción de los cinco controladores que componen el sistema nervioso central humano [Múgica y Nebot, 2003].

4.3 Selección de frecuencias para redes neuronales usando técnicas de selección de variables

Dentro de esta línea de trabajo se ha diseñado y probado un algoritmo para la construcción incremental de *FNNs* mediante la selección de datos de entrada que son usados como pesos (frecuencias) en las neuronas ocultas de la red [Romero y Alquézar, 2002a]. Actualmente se está adaptando el método para ser aplicado en una red neuronal heterogénea usando técnicas de selección de variables. Este trabajo nace de la fusión de tres conceptos: la neurona heterogénea [Belanche, 2000], la técnica *SAOCIF* [Romero, 2000] y los algoritmos de selección de variables [Molina *et al.*, 2002].

4.4 Cooperación de expertos locales para *FIR*

En el contexto de la metodología *FIR*, se ha desarrollado una estrategia para razonar en presencia de incertidumbre. Esta metodología permite realizar predicciones del sistema a través de dos esquemas diferentes. El primero corresponde a un esquema de predicción basada exclusivamente en reglas patrón, mientras que el segundo corresponde a un esquema difuso tipo *Sugeno* puro [Múgica *et al.*, 2003].

4.5 Construcción incremental y entrenamiento de redes neuronales recurrentes híbridas

A partir de los resultados de los trabajos experimentales con redes recurrentes *LSTM* [Ribeiro y Alquézar, 2001, 2002a], se han desarrollado las herramientas para la construcción incremental de redes neuronales recurrentes híbridas para el procesamiento

de secuencias temporales [Ribeiro y Alquézar, 2002b]. Este trabajo ha contado con la colaboración del profesor Juan Antonio Pérez Ortiz, de la Universidad de Alicante.

5 Trabajos iniciados recientemente

5.1 Funciones de similitud y disimilitud

Una línea de trabajo recién iniciada se centra en la formalización axiomática de los conceptos duales de “función de similitud” y “función de disimilitud”, estudio de medidas concretas y su posible aplicación al aprendizaje automático. Sobre esta base se ha estudiado el efecto sobre sus propiedades de la aplicación de transformaciones [Orozco y Belanche, 2004].

5.2 Modelado de cuencas fluviales

Otro de los trabajos recién iniciados en *SOCO* concierne el análisis de datos referentes al estatus ecológico de ríos de la cuenca mediterránea, parte del proyecto europeo STREAMES (www.streames.org). El objetivo último es el desarrollo de modelos del comportamiento funcional de las cuencas fluviales. Aplicaremos el *GTM*, en distintas variantes, para el agrupamiento, regeneración y visualización de datos. Resultados preliminares han sido presentados en [Vellido *et al.*, 2004; Vicente *et al.*, 2004]

6 Selección de publicaciones del grupo *SOCO* (últimos 7 años)

A continuación se listan algunas de las publicaciones de los últimos años del grupo, ordenadas por antigüedad decreciente. Una lista más completa se puede encontrar en la página <http://www.lsi.upc.es/dept/investigacion/sectia/soco/>

Libros: Capítulos

- Lisboa, P.J.G., Vellido, A. y Wong, H.: Outstanding issues for clinical decision support with neural networks. In: Malmgren, H.; Borga, M.; Niklasson, L. (eds.) *Artificial Neural Networks in Medicine and Biology*. Springer, London (2000) 63-71.
- Vellido, A., Lisboa, P.J.G. y Meehan, K.: Characterizing and segmenting the business-to-consumer e-commerce market using neural networks. In: Lisboa, P.J.G., Edisbury, B. and Vellido, A. (eds.) *Business Applications of Neural Networks*. World Scientific, Singapore (2000) 29-54.
- Múgica, F. y Nebot, A.: Simulated Annealing for Automated Definition of Fuzzy Sets in Human Central Nervous System Modeling. In: Sanfeliu, A. and Ruiz-Shulcloper, J. (eds.) *Lecture Notes in Computer Science*, Vol. 2905. Springer-Verlag, Berlin Heidelberg New York (2003) 545-553.

- Nebot, A.: Automatic Tuning of Fuzzy Partitions in Inductive Reasoning. In: Sanfeliu, A. and Ruiz-Shulcloper, J. (eds.) *Lecture Notes in Computer Science*, Vol. 2905. Springer-Verlag, Berlin Heidelberg New York (2003) 554-562.
- Carreras X., Márquez L., y Romero E.: Máquinas de Vectores Soporte. In: Orallo, J.H., Ramírez Quintana, M.J., Ferri Ramírez, C. (eds.) *Introducción a la Minería de Datos*. Pearson (2004) 353-382.

Revistas internacionales

- Carvajal, R. y Nebot, A.: Growth Model for White Shrimp in Semi-intensive Farming using Inductive Reasoning Methodology. *Computers and Electronics in Agriculture*, 19 (1998) 187-210.
- Lisboa, P.J.G., Kirby, S.J.P., Vellido, A., Lee, Y.Y.B. y El-Deredy, W.: Assessment of Statistical and Neural Network Methods in NMR Spectral Classification and Metabolite Selection. *Nuclear Magnetic Resonance in Biomedicine*, 11 (1998) 225-234.
- Nebot, A., Cellier, F.E. y Vallverdú, M.: Mixed Quantitative/Qualitative Modeling and Simulation of the Cardiovascular System. *Computer Methods and Programs in Biomedicine*, 55 (1998) 127-155.
- Belanche, Ll., Valdés, J.J., Comas, J., Roda, I.R. y Poch, M.: Towards a Model of the Input-Output Behaviour of Wastewater Treatment Plants Using Soft Computing Techniques. *Environmental Modeling and Software*, 14 (1999) 409-419.
- Jensen, E.W., Nebot, A., Caminal, P. y Henneberg, S.W.: Fuzzy Inductive Reasoning Applied to Identify a Causal Relationship among Haemodynamic Parameters, Auditory Evoked Potentials and Isoflurane. *British Journal of Anaesthesia*, 82 (1999) 25-32.
- Nebot, A., Cellier, F.E., y Múgica, F.: Simulation of Heat and Humidity Budget of Biosphere 2 without its Air Conditioning. *Ecological Engineering Journal*, 13 (1999) 333-356.
- Vellido, A., Lisboa, P.J.G. y Meehan, K.: Segmentation of the on-line shopping market using neural networks. *Expert Systems with Applications*, 17 (1999) 303-314.
- Vellido, A., Lisboa, P.J.G. y Vaughan, J.: Neural networks in business: a survey of applications (1992-1998). *Expert Systems with Applications*, 17 (1999) 51-70.
- Belanche, Ll., Valdés, J.J., Comas, J., Roda, I.R. y Poch, M.: Prediction of the bulking phenomenon in Wastewater Treatment Plants. *Artificial Intelligence in Engineering*, 14 (2000) 307-317.
- Lisboa, P.J.G., Vellido, A., y Wong, H.: Bias reduction in skewed binary classification with Bayesian neural networks. *Neural Networks*, 13 (2000) 407-410.
- Valdés, J.J., Belanche, Ll. y Alquézar, R.: Fuzzy Heterogeneous Neurons for Imprecise Classification Problems. *International Journal of Intelligent Systems*, 15 (2000) 265-276.
- Vellido, A., Lisboa, P.J.G. y Meehan, K.: Quantitative characterization and prediction of on-line purchasing behaviour: a latent variable approach. *International Journal of Electronic Commerce*, 4 (2000a) 83-104.
- Vellido, A., Lisboa, P.J.G. y Meehan, K.: The Generative Topographic Mapping as a principled model for data visualization and market segmentation: an electronic commerce case study. *International Journal of Computers Systems and Signals*, 1 (2000b) 119-138.
- Vellido, A., Lisboa, P.J.G. y Meehan, K.: A systematic quantitative methodology for characterizing the business-to-consumer e-commerce market. *SIGBIO Newsletter*, 20 (2000c)
- Velázquez-Lerma, S.A. y Alquézar, R.: Solving Partially Observable Markov Decision Processes by Optimization Neural Networks. *International Journal of Knowledge-Based Intelligent Engineering Systems*, 5 (2001) 163-170.
- Vellido, A. y Lisboa, P.J.G.: An electronic commerce application of the Bayesian Framework for MLPs: the Effect of Marginalization and ARD. *Neural Computing & Applications*, 10 (2001)

- Selva, A., Mijares, T., Solans, R., Labrador, M., Romero, E., Sopena, J.M. y Vilardell, M.: The Neural Network as a Predictor of Cancer in Patients with Inflammatory Myopathies. *Arthritis and Rheumatism*, 46 (2002) 2547-2548.
- Serratos, F., Alquézar, R. y Sanfeliu, A.: Synthesis of function-described graphs and clustering of attributed graphs. *International Journal of Pattern Recognition and Artificial Intelligence*, 16 (2002) 621-655.
- Armada, I., Pol, E., Sopena J.M., y Romero, E.: Actividad de ocio turístico y personas mayores: Análisis de diferencias psicosociales entre viajeros y no viajeros. *Encuentros en la Psicología Social*, 1 (2003) 33-39.
- Gómez, P., Nebot, A., Ribeiro, S., Alquézar, R., Múgica, F. y Wotawa, F.: Local maximum ozone concentration prediction using soft computing methodologies. *Systems Analysis Modelling Simulation*, 43 (2003) 1011-1031.
- Nebot, A., Múgica, F., Cellier, F.E. y Vallverdú, M.: Modeling and Simulation of the Central Nervous System Control with Generic Fuzzy Models. *Simulation: Transactions of The Society for Modeling and Simulation*, 79 (2003) 648-669.
- Selva, A., Romero, E., Sopena, J.M., Mijares, T., Solans, R., Labrador, M. y Vilardell M.: Reply to Roland Linder et al. *Arthritis and Rheumatism* 48 (2003) 1169-1170.
- Serratos, F., Alquézar, R. y Sanfeliu, A.: Function-described graphs for modelling objects represented by sets of attributed graphs. *Pattern Recognition*, 36 (2003) 781-798.
- Vellido, A., El-Deredy, W. y Lisboa, P.J.G.: Selective Smoothing of the Generative Topographic Mapping. *IEEE Transactions on Neural Networks*, 14 (2003) 847-852.
- Cellier, F.E., López, J., Nebot, A. y Cembrano, G.: Confidence Measures for Predictions in Fuzzy Inductive Reasoning. *International Journal of General Systems* (2004) Accepted for publication.
- Mugica, F. y Nebot, A.: Causal Relevancy: A New Concept to Improve the Prediction Accuracy of Dynamical Systems using Inductive Reasoning. *International Journal of General Systems* (2004) Accepted for publication.
- Nebot, A., Cellier, F.E. y Carvajal, R.: Fuzzy Inductive Reasoning for Variable Selection Analysis and Modeling of Biological Systems. *International Journal of General Systems* (2004) Accepted for publication.
- Romero, E., Marquez, L., y Carreras, X.: Margin Maximization with Feed-forward Neural Networks: A Comparative Study with Support Vector Machines and AdaBoost. *Neurocomputing*, 57 (2004) 313-344.
- Sopena, J.M. y Romero, E.: Redes Neuronales y Métodos Estadísticos Clásicos en el Diagnóstico Médico: La Importancia de las Variables Irrelevantes. *Medicina Clínica*, 122 (2004) Editorial.

Conferencias

- Sopena, J.M., Romero, E. y Alquézar, R.: Neural networks with periodic and monotonic activation functions: a comparative study in classification problems. *ICANN'99: Int. Conference on Artificial Neural Networks* (1999) 323-328.
- Ribeiro, S. y Alquézar, R.: A comparative study on a signal forecasting task applying Long-Short Term Memory (LSTM) recurrent neural networks. In: *Proc. of the SIARP'01, VI Simposio Iberoamericano de Reconocimiento de Patrones* (2001)113-119.
- Molina, L.C., Belanche, Ll., Nebot, A.: Feature selection algorithms: a survey and experimental evaluation. In: *Proc. of the IEEE Conference on Data Mining, Maebashi, Japón* (2002).
- Ribeiro, S. y Alquézar, R.: Local maximum ozone concentration prediction using LSTM recurrent neural networks. In: *Proc. of the EIS'2002, 3rd Int. NAISO Symposium on Engineering of Intelligent Systems*, 100030-04-SR-091 (2002a).

- Ribeiro, S. y Alquézar, R.: Incremental construction of LSTM recurrent neural network. In: Proc. CIARP'02, VII Congreso Iberoamericano de Reconocimiento de Patrones (2002b) 171-184.
- Romero, E. y Alquézar, R.: A New Incremental Method for Function Approximation using Feed-forward Neural Networks. IJCNN'02: International Joint Conference on Neural Networks (2002a) 1968-1973.
- Romero, E. y Alquézar, R.: Maximizing the Margin with Feed-forward Neural Networks. IJCNN'02: International Joint Conference on Neural Networks (2002b) 743-748.
- Múgica, F., Nebot, A. y Gómez, P.: Dealing with Uncertainty in Fuzzy Inductive Reasoning Methodology. UAI'03: 19th Conference on Uncertainty in Artificial Intelligence, Acapulco, México (2003) 427-434.
- Romero, E., Sopena, J.M., Navarrete, G. y Alquézar, R.: Feature Selection Forcing Overtraining May Help to Improve Performance. IJCNN'03: International Joint Conference on Neural Networks (2003) 2181-2186.
- Orozco, J., Belanche, L.I.: On Aggregation Operators of Transitivity Similarity. FUZZ-IEEE 2004, (2004) aceptado para su publicación.
- Romero, E., Carreras, X., Márquez, L.: Exploiting Diversity of Margin-based Classifiers. IJCNN'04: International Joint Conference on Neural Networks, (2004) aceptado para su publicación.
- Vicente, D., Vellido, A., Martí, E., Comas, J., and Rodríguez-Roda, I.: Exploration of the ecological status of Mediterranean rivers: Clustering, visualizing and reconstructing streams data using Generative Topographic Mapping. In W.I.T. Transactions on Information and Communication Technologies, Vol.30 (2004).

Tesis doctorales en curso

- Jesús Acosta, "Aprendizaje de particiones difusas para razonamiento inductivo", Dirigida por A. Nebot (UPC) y codirigida por J.M. Fuertes (UPC).
- Antoni Escobet, "Generación de decisiones delante de incertidumbre", dirigida por A. Nebot (UPC) y codirigida por F. Cellier (Univ. of Arizona).
- Pilar Gómez, "Fuzzy systems for pollution modelling in urban areas", Dirigida por A. Nebot (UPC) y codirigida por F. Mugica (México).
- Luis Carlos Molina, "Feature selection and discretization: A study applied to petroleum data sets", dirigida por L.I. Belanche (UPC) y codirigida por A. Nebot (UPC).
- Jorge Orozco, "Contribuciones al estudio e integración de sistemas neuronales y difusos", dirigida por L.I. Belanche (UPC).
- Sabrine Ribeiro, "Incremental methods for prediction using recurrent neural networks", dirigida por R. Alquézar (UPC).
- Enrique Romero, "Learning with Feed-forward Neural Networks: Three Schemes to Deal with the Bias-Variance Trade-off", dirigida por R. Alquézar y codirigida por J.M. Sopena (UB).

Referencias generales

Por orden alfabético:

Grupo SOCO

- Acosta, J. Nebot, A., Villar, P. y Fuertes, J.M.: Automatic learning of fuzzy partitions in human Central Nervous System modeling using genetic algorithms. WMC'04: Western Simulation Multiconference. International Conference on Health Science Simulation, San Diego (2004) 90-95.
- Alquézar, R.: Symbolic and Connectionist Learning Techniques for Grammatical Inference. Tesis doctoral, Univ. Politècnica de Catalunya (1997).
- Alquézar, R. y Sanfeliú, A.: Representation and recognition of regular grammars by means of second-order recurrent neural networks. In: Mira, J. Cabestany, J., and Prieto, A. (eds.) Trends in Neural Computation. Proc. of the Int. Workshop on Artificial Neural Networks IWANN'93, Sitges, Spain. Springer-Verlag, LNCS-686 (1993) 143-148.
- Alquézar, R. y Sanfeliú, A.: Inference and recognition of regular grammars by training recurrent neural networks to learn the next-symbol prediction task. In: Casacuberta, F. and Sanfeliú, A. (eds.) Advances in Pattern Recognition and Applications. World Scientific, Singapore (1994a) 48-59.
- Alquézar, R. y Sanfeliú, A.: An algebraic framework to represent finite-state machines in single-layer recurrent neural networks. Neural Computation, 7 (1995) 931-949.
- Alquézar, R. y Sanfeliú, A.: A hybrid connectionist-symbolic approach to regular grammatical inference based on neural learning and hierarchical clustering. In: Carrasco, R.C. and Oncina, J. (eds.) Grammatical Inference and Applications. Proceedings 2nd Int. Colloquium on Grammatical Inference, ICGI'94, Alicante, Spain. Springer-Verlag, LNAI-862 (1994b) 203-211.
- Alquézar, R. y Sopena, J.M.: Effect of unbounded activation functions on learning performance of recurrent networks. Tech. Rep. IC-DT-9402, Institut de Cibernètica, UPC-CSIC, Barcelona (1994).
- Alquézar, R., Sanfeliú, A. y Sainz, M.: Experimental assessment of connectionist regular inference from positive and negative examples. Preprints VII Simposium Nacional de Reconocimiento de Formas y Análisis de Imágenes, Centre de Visió per Computador, UAB, Bellaterra, Barcelona, 1 (1997) 49-54.
- Alquézar, R., Cueva, J., Valdés, J.J., Nebot, A. y Caminal, P.: Learning a multi-subject model of the central nervous system control using neural networks. In Proc. of the EIS'98, Int. ICSC Symposium on Engineering of Intelligent Systems, Tenerife, Spain, 2 (1998) 206-212.
- Belanche, L.A. y Valdés, J.J.: Using Fuzzy Heterogeneous Neural Networks to Learn a Model of the Central Nervous System Control. In Proc. of the EUFIT'98, 6th European Congress on Intelligent Techniques and Soft Computing, Aachen, Alemania (1998) 1858-1862.
- Belanche, L.A., Valdés, J.J. y Alquézar, R.: Fuzzy Heterogeneous Neural Networks for Signal Forecasting. In Proc. of the 8th Int. Conference on Artificial Neural Networks, ICANN'98, Skövde, Suecia (1998) 1089-1094.
- Belanche, L.: A study in function optimization with the Breeder Genetic algorithm. Informe técnico LSI-99-36-R, UPC (1999).
- Belanche, L.: Heterogeneous Neural Networks: Theory and Applications. Tesis doctoral, Universitat Politècnica de Catalunya (2000).
- Cueva, J., Alquézar, R. y Nebot, A.: Experimental comparison of fuzzy and neural network techniques in learning models of the central nervous system control. In Proc. of the EUFIT'97, 5th European Congress on Intelligent Techniques and Soft Computing, Aachen, Alemania, Verlag Mainz, Vol.II (1997) 1014-1018.
- Escobet, A., Nebot, A. y Cellier, F.E.: Model Acceptability Measure for the Identification of Failures in Qualitative Fault Monitoring Systems'. In Proc. of the ESM99, 13th European Simulation Multiconference, Warsaw, Poland (1999a) 339-347.
- Escobet, A., Nebot, A. y Vallverdú, M.: A Qualitative Fault Monitoring System to Support Medical Decisions: An Application to the Central Nervous System. In Proc. of the SCB'99,

- International ICSC Symposium on Soft Computing in Biomedicine, Rochester, NY (1999b) 536-542.
- Escobet, A., Huber, R., Nebot, A. y Cellier, F.E.: Enhanced Equal Frequency Partition Method for the Identification of a Water Demand System. In Proc. of the AI, Simulation and Planning Conference (AIS'00), Tucson, Arizona (2000) 209-215.
- Múgica, F. y Nebot, A.: A Specialization of the k -Nearest Neighbor Classification Rule for the Prediction of Dynamical Systems using FIR. In Advances in Artificial Intelligence and Engineering Cybernetics - Volume III, InterSymp'96, 8th International Conference on Systems Research Informatics and Cybernetics, Baden-Baden, Alemania (1996) 130-136.
- Nebot, A., Medina, S. y Cellier, F.E.: The Causality Horizon: Limitations to Predictability Predictability of Behavior Using Fuzzy Inductive Reasoning. In Proc. of the ESM-ICQFN'94, Barcelona, Spain (1994) 492-496.
- Nebot, A. y Múgica, F.: Using Causal Relevancy for the Selection of Models within FIR Qualitative Modeling and Simulation Environment. In Advances in Artificial Intelligence and Engineering Cybernetics - Volume III, InterSymp'96, 8th International Conference on Systems Research Informatics and Cybernetics, Baden-Baden, Alemania (1996) 137-142.
- Negri, S. y Belanche, L.I.: Heterogeneous Kohonen Networks. In Proc. of the 6th International work-conference on Artificial and Natural Neural Networks (IWANN'01), Granada, Spain (2001) 243-252.
- Quartier, B. y Belanche, L.I.: Algorithmes d'entraînement local de réseaux de neurones RBF. Technical Report LSI-01-42-R, UPC (2001).
- Romero, E.: Function Approximation in Hilbert Spaces: A General Sequential Method and a Particular Implementation with Neural Network. Technical Report LSI-00-13-R, UPC (2000).
- Sanfeliu, A. y Alquézar, R.: Understanding neural networks for grammatical inference and recognition. In Advances in Structural and Syntactic Pattern Recognition, Proceedings SSPR'92, Fourth IAPR Int. Workshop on Structural and Syntactic Pattern Recognition, Bern, Switzerland; Bunke, H. (ed.), World Scientific, Singapore (1992) 75-98.
- Sopena, J.M. y Alquézar, R.: Improvement of learning in recurrent networks by substituting the sigmoid activation function. In ; Marinaro, M. y Morasso, P.G. (eds.) Proc. of the ICANN'94, 4th Int. Conference on Artificial Neural Networks, Sorrento, Italia, 1 (1994) 417-420.
- Valdés, J.J. y García, R.: A model for heterogeneous neurons and its use in configuring neural networks for classification problems. In Proc. of the Int. Work-conference on Artificial and Natural Neural Networks IWANN'97, Lanzarote, Spain. Biological and Artificial Computation: From Neuroscience to Technology, Springer, LNCS-1240 (1997) 19-26.
- Vellido, A., Olier, I., Martí, E., Comas, J. y Roda, I.R.: STREAMES project: Exploration of the ecological state of Mediterranean rivers using Generative Topographic Mapping. In BESAI 2004 - 4th ECAI Workshop on Binding Environmental Sciences and Artificial Intelligence. Accepted (2004).

Otras

- Bishop, C.M., Svensén, M. y Williams, C.K.I. GTM: The Generative Topographic Mapping. Neural Computation, 10 (1998) 215-234.
- de Kleer, J. y Williams, B.: Diagnosis with Behavioral Modes. In Proc. of the 11th Intl. Conf. on Artificial Intelligence, Detroit, Mich., USA (1989).
- Schmidhuber, J.: A fixed size storage $O(n^3)$ time complexity learning algorithm for fully recurrent continually running networks. Neural Computation, 4 (1992) 243-248.