

INTRODUCCION DE UN MECANISMO PARA LA MEJORA DE LA INTERPRETABILIDAD EN SISTEMAS DIFUSOS LINGÜÍSTICOS CON DEFUZZIFICACION ADAPTATIVA

A.A. Márquez F.A. Márquez A. Peregrín

Dpto. de Tecnologías de la Información. Universidad de Huelva 21071 – Huelva,
{amarquez, alfredo.marquez, peregrin}@dti.uhu.es,

Resumen

En este trabajo se propone un mecanismo para la mejora de la interpretabilidad en sistemas difusos lingüísticos con defuzzificación adaptativa. El uso de parámetros en el operador de defuzzificación introduce una serie de valores o pesos asociados indirectamente a cada regla, que mejoran su precisión pero aumentan la complejidad y por tanto perjudican a la interpretabilidad del sistema. En este trabajo se trata de reducir dicha complejidad pretendiendo obtener tanto las ventajas de una defuzzificación adaptativa en cuanto a precisión, como reducir la pérdida de interpretabilidad. Para tal fin, utilizaremos un modelo evolutivo multiobjetivo que proporcionará un conjunto de sistemas difusos lingüísticos, con el mejor equilibrio entre el número de reglas finales, el número de reglas con peso y la precisión del sistema.

Palabras Clave: Modelado Difuso Lingüístico, Interpretabilidad, Algoritmos Genéticos Multiobjetivo, Sistema de Defuzzificación Adaptativo, Selección de Reglas, Reglas Ponderadas.

1 INTRODUCCION

En la actualidad, el problema de encontrar un equilibrio entre interpretabilidad y precisión en Sistemas Basados en Reglas Difusas Lingüísticas, a pesar de la naturaleza intrínseca de la lógica difusa, ha dado lugar a un creciente interés en métodos que tienen en cuenta ambos aspectos [4, 5]. Por supuesto, lo ideal sería satisfacer ambos criterios en un grado alto, pero debido a que son características contradictorias, por lo general esto no es

posible. Un camino para conseguirlo es mejorar la precisión del sistema mientras se mantiene en lo posible la interpretabilidad a un nivel aceptable [4].

En este ámbito, el sistema de inferencia y los métodos de defuzzificación adaptativos han demostrado ser unos elementos capaces de mejorar fácilmente la precisión del sistema [1, 8, 15] debido a que buscan la mejor forma de inferir la contribución de cada regla de forma apropiada [8] y mejorar la cooperación entre las reglas [1]. Estos pueden también adaptar especialmente el comportamiento de los operadores difusos a la base de reglas (BR) o aprender los operadores difusos y la BR conjuntamente [15], obteniendo una sinergia positiva entre ambos elementos que permite al sistema alcanzar un nivel alto de precisión con la pérdida de interpretabilidad asociada al uso de operadores adaptativos con sus correspondientes parámetros.

Por otro lado, en la literatura se distinguen dos clases de aproximaciones en interpretabilidad [3, 17]:

1. Interpretabilidad Global (estructura del sistema): La complejidad del modelo (generalmente mediada como número de reglas, variables, etiquetas por variable, etc.)
2. Interpretabilidad Local (comprensión del modelo): Una medida semántica de la interpretabilidad (sistema de inferencia utilizado, método de defuzzificación, operador de conjunción, etc.)

Por tanto, la defuzzificación adaptativa [8], introducirá una pérdida de interpretabilidad en principio en el sentido local, con efectos en sentido global, en tanto los parámetros asociados a cada regla aumentan la estructura del sistema con la inclusión de un peso. Sería deseable pues, minimizar los elementos que aumenten la complejidad del sistema.

En este trabajo, se propone introducir, en primer lugar, un mecanismo que permita mejorar la interpretabilidad cuando se utilizan defuzzificación adaptativa. Dicho mecanismo se basará en dos índices que permitan

eliminar el peso de la regla, o bien la regla completa en el mismo proceso de aprendizaje evolutivo de la defuzzificación adaptativa.

En segundo lugar, se propone utilizar un Algoritmo Genético Multi-objetivo (AGMO) [7, 9] como modelo evolutivo para obtener modelos difusos lingüísticos precisos e interpretables utilizando la defuzzificación adaptativa con tres objetivos (minimizar la precisión, y minimizar los dos índices de interpretabilidad comentados anteriormente). De esta forma, se obtendrá un conjunto de soluciones con distinto equilibrio entre precisión e interpretabilidad [11, 12, 13, 14].

El trabajo se estructura de la siguiente forma: en la Sección 2 se repasa la construcción de la defuzzificación adaptativa utilizada, sus componentes y efectos. En la Sección 3, se presenta las bases del mecanismo de mejora de interpretabilidad propuesto. La Sección 4 presenta el modelo multi-objetivo, describiendo sus principales características y operadores genéticos considerados. La Sección 5 se dedica a un estudio experimental en el que se emplean las técnicas anteriormente descritas en dos problemas reales, para finalmente presentar las conclusiones del estudio desarrollado.

2 MÉTODOS DE DEFUZZIFICACIÓN ADAPTATIVOS EVOLUTIVOS

Existen varias tendencias en la literatura en el desarrollo de métodos de defuzzificación adaptativa. Éstos emplean uno o más parámetros en sus expresiones con el fin de modificar el comportamiento de la defuzzificación o, en la mayoría de los casos, mejorar la precisión del sistema de forma considerable.

Siguiendo los estudios desarrollados en [8], y debido a su fácil implementación y buen comportamiento, en el presente trabajo hemos considerado la utilización de un método de defuzzificación adaptativo basado en (1):

$$y_0 = \frac{\sum_i^N f(h_i) \cdot V_i}{\sum_i^N f(h_i)}, \quad (1)$$

donde h_i es el grado de emparejamiento, $f(h_i)$ es un término funcional del grado de emparejamiento y V_i es el máximo valor (MV_i) o el centro de gravedad (CG_i). Como puede observarse, se trata de la expresión de un método de defuzzificación que actúa en Modo B, es decir, primero se defuzzifica la aportación individual de la inferencia con cada regla y posteriormente se computa el resultado final, concretamente mediante una suma ponderada.

Concretamente, hemos considerado el uso de un término funcional tipo producto $f(h_i) = h_i \cdot \alpha_i$, donde α_i corresponde a un parámetro para cada regla R_i , $i=1$ a N , y como valor característico el Centro de Gravedad (CG), debido a su eficiencia computacional y obtener similares resultados a otro tipo de funciones [8]. La expresión de dicho método de defuzzificación adaptativa se muestra en (2),

$$y_0 = \frac{\sum_i^N h_i \cdot \alpha_i \cdot CG_i}{\sum_i^N h_i \cdot \alpha_i}, \quad (2)$$

Como se puede apreciar, la influencia del parámetro individual α se puede interpretar como una modulación de la influencia del grado de emparejamiento, potenciándolo o penalizándolo. Particularmente, el estudio zonal del efecto de α es el siguiente:

- $\alpha \cdot h_i$, $\alpha \in [1, \infty)$: potenciación de h_i ,
- $\alpha \cdot h_i$, $\alpha \in [0, 1]$: penalización de h_i .

Además, el funcional de tipo producto con un parámetro para cada regla tiene un efecto equivalente a utilizar reglas con pesos [6]. El empleo de pesos que acompañan a cada regla de la base de conocimiento consiste en añadir un valor adicional que tiene el significado de indicar qué grado de importancia debe tener esa regla en el proceso de inferencia. Generalmente el objetivo de esta ampliación es mejorar la precisión.

3 MECANISMO PARA LA MEJORA DE LA INTERPRETABILIDAD EN LOS SISTEMAS DIFUSOS LINGÜÍSTICOS CON DEFUZZIFICACIÓN ADAPTATIVA

En esta sección, proponemos introducir un mecanismo para mejorar la interpretabilidad en sistemas difusos lingüísticos con defuzzificación adaptativa, basado en una métrica.

Tal y como se ha indicado en la sección anterior, la utilización de métodos de defuzzificación adaptativos con funcional de tipo producto tiene un efecto equivalente al empleo de reglas con pesos [6]. Estos pesos en las reglas tienen un efecto negativo en la interpretabilidad [3, 17] ya que extienden la estructura del sistema. Por este motivo, y con objeto de reducir este efecto negativo, la métrica que se utilice debe tenerlo en cuenta. En este trabajo se ha optado por utilizar el número de reglas final como medida útil en este sentido, así como el número de reglas con peso, considerando que ambos valores son mejores

cuando menores sean. En las siguientes subsecciones se describirán con más detalle estas métricas.

3.1. NÚMERO DE REGLAS FINAL ($\#R_F$)

Esta métrica está basada en la idea de que aquellos pesos de las reglas próximos a 0 suponen una baja influencia de dicha regla y por tanto, podrían señalar una regla prescindible, continuando el sistema evolutivo con el aprendizaje del resto del sistema sin ella. Eliminar la regla en la fase de aprendizaje permite al modelo evolutivo ajustar el resto de pesos para eliminar dicha regla. Si se llevase a cabo la eliminación de las reglas con poco peso sólo al final del proceso, la precisión del sistema, en la mayoría de los casos se vería perjudicada por la naturaleza de los sistemas difusos lingüísticos.

Para aplicar dicho concepto los parámetros del funcional tipo producto actuarán en el rango [0,1] (véase expresión 2). La expresión de dicho índice será:

$$\#R_F = \#R - (\text{número de pesos próximos a 0}) \quad (3)$$

donde $\#R$ es el número de reglas iniciales del sistema.

3.2. NUMERO DE REGLAS CON PESO ($\#R_P$)

Esta métrica en cambio, se basa en la idea de que los valores de los pesos próximos a 1 son aquellos en los que la regla tiene importancia y podría utilizarse sin necesidad de ninguna ponderación o peso, y por tanto eliminar este valor reduciendo así la complejidad del sistema. Cuanto mayor sea el número de reglas ponderadas por un peso, mejor será la interpretabilidad del sistema. La expresión de dicho índice será:

$$\#R_P = \#R - (\text{número de pesos próximos a 1}) \quad (4)$$

3.3. INTRODUCCION DE LOS INDICES EN EL PROCESO DE APRENDIZAJE

Como se ha apuntado anteriormente, la introducción de dichos índices en el proceso de aprendizaje es un elemento importante, ya que el sistema debe readaptarse tanto a la eliminación de reglas con baja aportación, como a la de algunos de los pesos en las reglas más importantes.

El mecanismo que se utiliza en este trabajo es proporcionar al sistema unos umbrales fijos para ambos índices, los cuales son tenidos en cuenta en la evaluación del sistema difuso. Los pesos que superen un valor determinado próximo a 1, serán evaluados como 1, mientras que los valores del peso que desciendan por debajo de un umbral cercano a 0, serán reglas no consideradas.

4 MODELO MULTI OBJETIVO

En esta Sección se presenta el algoritmo multiobjetivo propuesto donde se introducirá el mecanismo de mejora de interpretabilidad en sistemas con defuzzificación adaptativa descrito en la Sección anterior.

Los objetivos que utiliza el modelo multiobjetivo son tres: minimizar los dos índices de interpretabilidad, $\#R_F$ y $\#R_P$ además de minimizar el habitual error cuadrático medio del sistema para la precisión. En esta ocasión se utiliza un modelo evolutivo basado en el conocido NSGA-II [10]. En las próximas subsecciones se expondrán los principales componentes de dicho algoritmo.

3.1. ESQUEMA DE CODIFICACIÓN Y POBLACIÓN INICIAL

En este trabajo se utiliza un esquema de codificación real, siendo m parámetros α_i , uno para cada R_i de la BR. Cada uno toma valores en el intervalo [0,1].

$$C = (\alpha_1, \dots, \alpha_m) \mid \alpha_i \in \{0, 1\}$$

La población inicial se obtiene de la siguiente forma: Un individuo de la población inicial tiene todos los genes inicialmente fijados a 1 con el fin de partir en el proceso evolutivo con todas las reglas sin peso. El resto de individuos de la población inicial se crean aleatoriamente.

4.2 OBJETIVOS Y UMBRALES DEL MECANISMO DE MEJORA

Tal y como hemos comentado anteriormente, en este algoritmo hemos usado estos objetivos a minimizar:

- El índice de interpretabilidad ($\#R_F$) que representa el número de reglas finales en el sistema.
- El índice de interpretabilidad ($\#R_P$) que representa el número de reglas que actúan con peso.
- El Error Cuadrático Medio (ECM) que mide la precisión del sistema:

$$ECM(S[i]) = \frac{\frac{1}{2} \sum_{k=1}^N (y_k - S[i](x_k))^2}{N} \quad (5)$$

donde $S[i]$ es el modelo difuso. Esta medida precisa del empleo de un conjunto de datos de evaluación constituido por N pares de datos numéricos $Z_k = (x_k, y_k)$, $k=1, \dots, N$, siendo x_k los valores de las variables de entrada e y_k los correspondientes a sus salidas asociadas.

En cuanto a los umbrales considerados para los índices $\#R_F$ y $\#R_P$, se han seleccionado los siguientes:

- Cuando el valor del parámetro es ≥ 0.9 se considera que dicho parámetro es 1 y por tanto dicha regla no tiene peso.

- Cuando el valor del parámetro sea ≤ 0.1 , evaluaremos dicho parámetro como 0 y dicha regla será eliminada.

Ambos umbrales han sido escogidos empíricamente realizando distintas pruebas.

4.3 ALGORITMO GENÉTICO MULTI OBJETIVO

El algoritmo genético elegido, basado en el NSGA-II [10], es uno de los AGMO más conocidos y utilizados en la literatura para la resolución de problemas multiobjetivos. La población descendiente se genera desde la población actual a través de la selección, el cruce y la mutación. La siguiente generación se construye pues desde la población actual y la descendiente hasta que se cumpla la condición de parada, en este trabajo, por número de evaluaciones.

El algoritmo NSGA-II posee dos características que lo hacen uno de los principales y más importantes AGMO: Una es la asignación del *fitness* basada en el *ranking* del pareto y en el operador de *crowding*, y la otra es el procedimiento de actualización de cada generación a través del elitismo.

5 ESTUDIO EXPERIMENTAL

Para analizar el funcionamiento del método propuesto, se han escogido dos problemas reales con diferentes complejidades (diferente número de variables y datos). Dichos problemas están disponibles en <http://www.keel.es/> [2]:

- La estimación del cálculo de los costes óptimos teóricos de la línea de media tensión en núcleos urbano (ELE).
- Predicción de la edad del Abalone (ABA).

En ambos casos, se ha empleado el algoritmo de Wang y Mendel [16] para obtener la BR inicial. Una vez generado, el algoritmo propuesto fue aplicado. Los Métodos considerados para el estudio experimental fueron:

- D método con defuzzificación adaptativa utilizando un algoritmo genético mono-objetivo y considerando sólo precisión como único objetivo [8].
- D_1 es el que utiliza la defuzzificación adaptativa con la mejora propuesta en la interpretabilidad y aprendizaje multiobjetivo considerando tres los objetivos, minimizar ECM $\#R_F$ y $\#R_p$.

Se ha considerado un modelo experimental 5-fold cross-validation. Los métodos han sido ejecutados 6 veces, mostrándose la media de las 30 ejecuciones.

Las particiones lingüísticas consideradas están formadas por términos lingüísticos con forma triangular, 5 para el problema eléctrico y 3 para el Abalone. Como operador de conjunción e implicación se ha utilizado la t-norma del mínimo. Los valores de los parámetros utilizados en todos los experimentos presentados son: 200 de longitud de la población, 200 de longitud de la descendiente, 1000 generaciones y 0.2 como probabilidad de mutación por cromosoma.

5.1 RESULTADOS Y ANÁLISIS

En el caso de los métodos con enfoque multi-objetivo, las medias han sido obtenidas considerando la solución más precisa de cada pareto obtenido. Los algoritmos propuestos se han comparado con varios métodos guiados por un único objetivo.

La Tabla 1 muestran, a nivel de referencia, los ECMs de los conjuntos de entrenamiento y prueba en cada aplicación, donde $\#R$ representa el número de reglas, ECM_{ent} y ECM_{pru} representan el error obtenido en entrenamiento y en prueba respectivamente. Este método obtiene la BR inicial sobre la que D , y D_1 introducirán sus pesos. WM toma para el índice de $\#R_p$ el más alto nivel que representa que todas las reglas no contienen pesos.

Tabla 1: ECMs de referencia

Problema	#R	ECM_{ent}	ECM_{pru}
ELE	65	56136	56359
ABA	68	8.407	8.422

Los resultados obtenidos con los métodos de defuzzificación adaptativos se muestran en la Tablas 2 donde adicionalmente se añade dos columnas con los valores de $\#R_F$ (representa el número de reglas finales) y $\#R_p$ (el número de reglas con peso) como índices de interpretabilidad, y dos columnas denominadas t-test que son los resultados de aplicar un test t-student (con 95% de confianza) al mejor resultado medio de la correspondiente columna comparado 1 a 1 con el resto de resultados medios para cada problema. La interpretación de ellas es:

- * Indica el resultado con la mejor media.
- + Indica un comportamiento significativamente peor que el mejor.
- = Indica un comportamiento estadísticamente sin diferencias significativas con el mejor.

Observando los resultados presentados en la Tabla 2 podemos destacar:

- El modelo que considera las medidas de interpretabilidad obtiene el mejor comportamiento en cuanto a interpretabilidad respecto al que no, D , en ambos problemas. D_1 mejora tanto en $\#R_F$ como en $\#R_p$, con descensos significativos en el número de

reglas, y siendo particularmente interesante en ambos problemas que sólo aproximadamente la mitad de las reglas finales tienen pesos.

- En cuanto a la precisión, en el problema del ABA se obtiene similares niveles de precisión tanto en entrenamiento como prueba, mientras que en el problema ELE la precisión es algo mayor en el modelo mono-objetivo orientado a precisión, señalando por tanto la necesidad de mejoras en el modelo multiobjetivo para alcanzar los mismos niveles de precisión.

En la Figura 1 se muestra la superficie de la frontera del pareto del algoritmo D_1 del problema ABA. Se puede observar que es bastante extenso y que no existe sobreaprendizaje en los resultados obtenidos con el modelo propuesto.

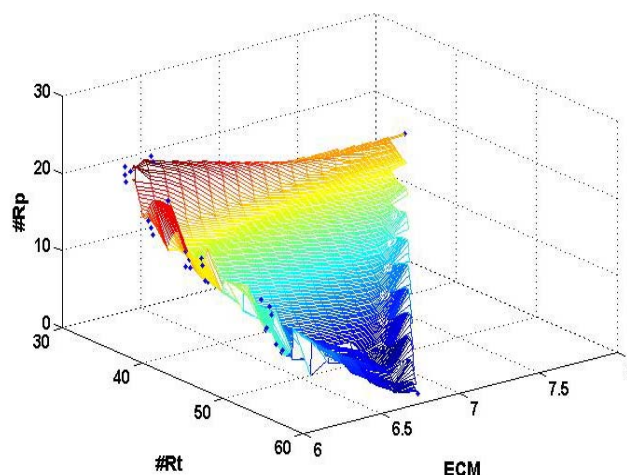


Figura 1: Frontera de pareto obtenida en el problema ABA.

Tabla 2: Resultados obtenidos en los dos problemas

Problema	Método	ECM _{ent}	t-test	ECM _{pru}	t-test	#R _F	t-test	#R _P	t-test
ELE	D	32791	*	35362	*	65	+	65	+
	D_1	34921	+	37695	+	45,97	*	19,63	*
ABA	D	4.823	=	4.826	*	68	+	68	+
	D_1	4,918	=	4,924	=	31,57	*	17,70	*

6 CONCLUSIONES

La defuzzificación adaptativa que utiliza factores de ponderación en el grado de emparejamiento a través de un producto es un mecanismo sencillo para mejorar la precisión de los modelos difusos lingüísticos, el cual, tiene como desventaja el aumento de la complejidad del sistema, resultado del efecto de añadir distintos factores o pesos en cada una de las reglas de la base de conocimiento. Para reducir este efecto, en este trabajo introducimos un mecanismo para reducir complejidad basado en utilizar umbrales sobre los citados pesos, de modo que niveles muy altos eliminan el peso de esa regla y niveles muy bajos eliminan la regla por completo. Aprendiendo mediante algoritmos evolutivos multiobjetivo los parámetros tratando de obtener la mayor precisión, el menor número de reglas y el menor número de reglas con pesos, se obtienen resultados interesantes por la reducción en la complejidad, manteniendo un buen nivel de precisión, como se ha visto experimentalmente.

No obstante, observando la Figura 1, consideramos ampliar el estudio tratando de centrar el algoritmo evolutivo multiobjetivo en la zona más interesante de la frontera del pareto. Además, como trabajo futuro se desea tener en cuenta dos aspectos:

- La elección de los umbrales empleados podría ser más eficaz si se eligieran en base al problema y su dominio.
- Mejorar la propuesta de interpretabilidad teniendo en cuenta otros índices como son la inconsistencia, redundancia y similaridad de las reglas.

Agradecimientos

Este trabajo está soportado por el proyecto TIN2008-06681-C06-06 del Ministerio de Educación y Ciencia y el proyecto P07-TIC-03179 de la Consejería de Innovación Ciencia y Empresa de la Junta de Andalucía.

Referencias

- [1] J. Alcalá-Fdez, F. Herrera, F. Márquez, A. Peregrín. Increasing Fuzzy Rules Cooperation Based On Evolutionary Adaptive Inference Systems. *International Journal of Intelligent Systems* 22:9, Pág. 1035-1064, 2007.
- [2] J. Alcalá-Fdez, L. Sánchez, S. García, M.J. del Jesús, S. Ventura, J.M. Garrell, J. Otero, C. Romero, J. Bacardit, V.M. Rivas, J.C. Fernández, and F. Herrera. Keel: A software tool to assess evolutionary

- algorithms for data mining problems. *Soft Computing*, 13:3, Pág. 307–318, 2009.
- [3] J.M. Alonso, L. Magdalena. A Conceptual Framework for Understanding Fuzzy Systems. *Proceedings of IFSA World Congress (IFSA-EUSFLAT'09)* Pág. 119-124, Lisbon, Portugal, 2009
- [4] J. Casillas, O. Cordon, F. Herrera, L. Magdalena(Eds). Interpretability issues in fuzzy modeling, volume 128 of Studies in Fuzziness and Soft Computing. *Springer*, 2003.
- [5] J. Casillas, O. Cordon, F. Herrera, L. Magdalena(Eds). Accuracy improvements in linguistic fuzzy modeling. *Springer-Verlag*, 2003.
- [6] J.S. Cho, D.J. Park. Novel fuzzy logic control based on weighting of partially inconsistent rules using neural network. *Journal of Intelligent and Fuzzy Systems*, 8, Pág. 99-100, 2000
- [7] C. A. Coe llo, D. A. Van Veldhuizen, and G. B. Lamont, editors. Evolutionary algorithms for solving multi-objective problems. *Kluwer Academic Publishers*, 2002.
- [8] O. Cordon , F. Herrera , F. A. Márquez, A. Peregrín. A study on the evolutionary adaptive defuzzification methods in fuzzy modelling. *International Journal of Hybrid Intelligent Systems*. 1:1. Pág. 36–48, 2004.
- [9] K. Deb. Multi-objective optimization using evolutionary algorithms. *John Wiley & Sons*, NY, USA, 2001.
- [10] K. Deb, A. Pratap, S. Agarwal, T. Meyarivan. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Trans. on Evolutionary Computation* 6:2, Pág.182-197, 2002.
- [11] M J Gacto, R Alcalá, F Herrera. Adaptation and application of multi-objective evolutionary algorithms for rule reduction and parameter tuning of fuzzy rule-based systems. *Soft Computing* 13. Pág. 419-436, 2009.
- [12] H. Ishibuchi, T. Yamamoto. Fuzzy rule selection by multi-objective genetic local search algorithms and rule evaluation measures in data mining. *Fuzzy Sets Syst* 141:1. Pág. 59–88, 2004.
- [13] A.A. Márquez, F.A. Márquez, A. Peregrín. Cooperation between the Inference System and the Rule Base by using Multi-objective Genetic Algorithms. *Proceedings of Third International of Hybrid Artificial Intelligence System,(HAIS'08)*. Pág. 739-746 Burgos, Spain, 2008.
- [14] A.A. Márquez, F.A. Márquez, A. Peregrín. Rule Base and Inference System Cooperative Learning of Mamdani Fuzzy Systems with Multiobjective Genetic Algorithms. *Proceedings of IFSA World Congress (IFSA-EUSFLAT'09)* Pág. 1045-1050, Lisbon, Portugal, 2009.
- [15] F.A. Márquez, A. Peregrín y F. Herrera. Cooperative Evolutionary Learning of Fuzzy Rules and Parametric Aggregation Connectors for Mamdani Linguistic Fuzzy Systems. *IEEE Transactions on Fuzzy Systems*. 15:6 Pág. 1162-1178, 2007.
- [16] L.X. Wang, J.M. Mendel. Generating fuzzy rules by learning from examples. *IEEE Trans. on Systems, Man, and Cybernetics* 22:6, Pág.1414-1427, 1992.
- [17] M. Zhou and J. Q. Gan. Low-level interpretability and high-level interpretability: a unified view of data-driven interpretable fuzzy system modelling. *Fuzzy Sets and Systems*, 159:23. Pág. 3091–3131, 2008.