SISTEMAS BASADOS EN REGLAS DIFUSAS CON DATOS IMPRECISOS: NUEVOS RETOS

Luciano Sánchez[1] **Inés Couso**[2]

[1] Departamento de Informática, Universidad de Oviedo, luciano@uniovi.es
[2] Departamento de Estadística e I.O. y D.M., Universidad de Oviedo, couso@uniovi.es

Resumen

En las bases de conocimiento difusas, los conjuntos borrosos representan conceptos imprecisos, contenidos en los antecedentes y consecuentes de las reglas. Un sistema basado en reglas difusas (SBRD) contiene, además de la base de conocimiento, un motor de inferencia, con interfaces que le permiten admitir y producir datos numéricos. En este trabajo discutiremos cómo alterar este motor de inferencia de forma que el SBRD pueda ser aplicado al procesamiento de datos borrosos.

Además, analizaremos los retos que plantea el uso de este nuevo tipo de información: cómo representar datos vagos mediante conjuntos borrosos, cómo preprocesar esta información, qué esquemas de razonamiento son más adecuados, qué tipos de algoritmos permiten dotar a los SBRD de capacidad de aprendizaje a partir de datos de baja calidad y cómo validar los resultados de estos sistemas.

Palabras Clave: Sistemas Basados en Reglas Difusas, Datos Imprecisos.

1. INTRODUCCIÓN

Las bases de conocimiento borrosas contienen reglas "SI-ENTONCES", cuyos antecedentes y consecuentes contienen sentencias basadas en la lógica borrosa. Un Sistema Basado en Reglas Difusas (SBRD) combina una base de conocimiento y un motor de inferencia. Este último contiene un sistema de inferencia, e interfaces de fuzzificación y defuzzificación [6]. Tales interfaces convierten datos numéricos en información borrosa, compatible con las reglas que componen la base de conocimiento, y transforman los resultados borrosos de la base de conocimiento en valores nítidos. Como es bien conocido, esta estructura permite que los SBRDs puedan aproximar una amplia familia de funciones con precisión arbitraria [5]. Comparados con otros aproximadores universales, los SBRDs tienen un valor adicional, porque balancean la precisión con la interpretabilidad linguistica [4], explotando el potencial de la lógica borrosa como un esquema de representación y cálculo de nociones inciertas o vagas.

Paradójicamente, los SBRDs, pese a estar basados en la lógica borrosa, no se aplican con frecuencia a datos imprecisos. La práctica común en SBRDs consiste en asociar conjuntos borrosas a palabras y usarlos para modelar afirmaciones vagas, pero hacer que los modelos que dependen de esas afirmaciones usen datos nítidos, y produzcan resultados nítidos. En este trabajo estudiaremos cómo superar esta limitación, y expondremos los nuevos retos que plantea el uso de datos de baja calidad en el contexto de los SBRDs.

La estructura de este documento es la siguiente: En la sección 2 se discute la representación de un dato vago mediante conjuntos borrosos, y se establecen los beneficios esperados del uso de datos imprecisos en problemas de clasificación y regresión. En la sección 3 se consideran algunas herramientas de análisis gráfico y preprocesamiento que permiten mejorar la calidad de la información imprecisa. En la sección 4 se estudia el diseño del motor de inferencia de un SBRDs que procesa datos vagos. En la sección 5 se recogen algunas medidas de precisión de tales sistemas. En la sección 6, el problema de validar los resultados de esos sistemas de forma comparada. Por último, en la sección 7 (conclusiones) se resumen todos los problemas abiertos.

2. REPRESENTACION DE DATOS DE BAJA CALIDAD

Los modelos basados en datos imprecisos no son exclusivos de la literatura relacionada con los sistemas borrosos, y constituyen un campo actual de investigación en varias disciplinas. Por ejemplo, en [14] se proponen ocho fuentes de información que pueden ca-

racterizarse mediante intervalos: informes de pasa-no llega, dígitos significativos, mediciones intermitentes, fallos en detección, datos censurados, datos agrupados, datos perdidos y total ignorancia.

Estos modelos tienen una fuerte relación con algunas tendencias recientes en estadística borrosa, donde el tratamiento formal de los datos vagos está basado en el concepto de Variable Aleatoria Borrosa (VAB). Una VAB es una función, que asocia cada resultado de un experimento aleatorio con un subconjunto borroso de la recta real. En general, puede considerarse que estos subconjuntos borrosos modelan la imprecisión en la observación del evento, mientras que la aleatoriedad en el experimento modela la variabilidad de los resultados [3, 10].

2.1. Definiciones de VAB

Hay diferentes definiciones de VAB, y la elección de una de ellas condiciona el uso conjunto de la lógica y la estadística borrosas necesario para aplicar SBRDs a datos vagos, por lo que a continuación se enumeran los modelos existentes en la literatura y se propone la adopción de uno de ellos.

2.1.1. Interpretación clásica

En [23], Puri y Ralescu consideran que las observaciones de algunos experimentos aleatorios no producen resultados numéricos, sino términos lingüísticos imprecisos. Si la VAB tiene un número finito de imágenes, pueden asignárseles valores de probabilidad. Por ejemplo, el resultado es "alto" con probabilidad 0.5, "medio" con probabilidad 0.25 y "bajo" con probabilidad 0.25, donde "alto", "medio" y "bajo" forman una partición borrosa.

Observe que, aunque esta definición considera que los resultados del experimento tienen un significado lingüístico, no necesariamente es compatible con el uso de estos términos en lógica borrosa. Por ejemplo, en el modelo anterior, la probabilidad de que el resultado sea el conjunto borroso "alto o medio" (entendiendo la disyunción como combinación de las pertenencias de ambas etiquetas con una t-conorma) es cero, porque el conjunto borroso asociado a esa expresión lógica no coincide exactamente con ninguna de las imágenes de la variable.

2.1.2. Modelo de Segundo Orden

Según [10], que a su vez extiende la definición de [20], una variable aleatoria borrosa representa un conocimiento impreciso acerca de una variable aleatoria clásica, que es llamada "variable aleatoria original". Es decir: supongamos que el resultado del experimento aleatorio es el valor ω, y que la observación del experimento aleatorio es un conjunto borroso con pertenencia $X(\omega)$. Entonces, la pertenencia de un punto x al conjunto borroso $X(\omega)$ representa el grado de posibilidad de la afirmación "La imagen de ω es x". O lo que es lo mismo, una VAB es un conjunto borroso que asocia una pertenencia a cada una de las posibles variables aleatorias clásicas. Dicha pertenencia indica el grado de certeza que se tiene acerca de que esta variable sea la original.

2.1.3. Modelo de Primer Orden

De acuerdo con [2], se supone que, si el resultado del experimento aleatorio es ω, la probabilidad de percibir el valor x es $p(x|\omega)$, y que esta segunda probabilidad no es completamente conocida: sólo se sabe que está dominada por la posibilidad $\Pi(x|\omega)$. En este modelo, la función de pertenencia $X(\omega)$ de la VAB es la posibilidad $\Pi(x|\omega)$. En otras palabras, una variable aleatoria borrosa se corresponde con una familia de distribuciones de probabilidad. Esto, como se indica en la figura 1, es una generalización directa de los modelos estocásticos usados habitualmente en ingeniería para caracterizar el ruido en las observaciones.

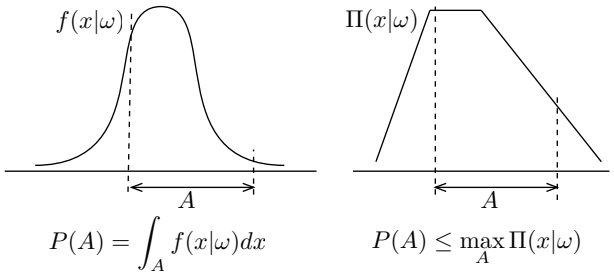

Figura 1: Modelos estocástico y posibilístico de primer orden del error de observación. **Izquierda:** Modelo estocástico de una observación con ruido. Se conoce de forma precisa la probabilidad de percibir un valor contenido en el conjunto A, cuando el resultado del experimento es ω. **Derecha:** modelo posibilístico de una observación vaga: se conoce un intervalo de valores que contiene a $P(A)$.

2.1.4. Discusión

En general, el modelo clásico no siempre es válido para representar datos vagos en el contexto de los SBRDs, ya que es problemático conciliar el uso de términos lingüísticos desde el punto de vista de la lógica borrosa con esta definición, como se ha visto en el ejemplo anterior. Por el contrario, tanto el modelo de primer orden como el modelo de segundo orden son adecuados. El tratamiento del modelo de primer orden tiene una implementación considerablemente más sencilla, y por

esta razón es preferible. Adicionalmente, de acuerdo con este modelo, los α-cortes de los conjuntos borrosos son intervalos de confianza del valor percibido de la variable, lo que permite representar mediante conjuntos borrosos algunos tipos de datos de baja calidad que tradicionalmente no se consideraban borrosos. Por ejemplo, en [24] se utilizaron conjuntos borrosos para reconciliar diferentes opiniones acerca de un valor, y en [22] se usó esta propiedad para asimilar las coordenadas proporcionadas por un GPS a conjuntos borrosos.

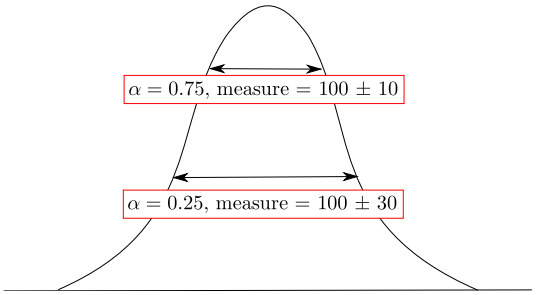

Figura 2: Visión alternativa del modelo de primer orden: conjunto borroso como familia de intervalos. Cada α-corte contiene al resultado del experimento con probabilidad $\geq 1 - \alpha$.

2.2. Beneficios esperados del uso de datos imprecisos

Una utilidad obvia de la extensión de los SBRDs a datos de baja calidad es la posibilidad de solucionar problemas en que no se dispone de mejor información. No obstante, esta técnica también proporciona ventajas en algunos problemas con datos nítidos. Por ejemplo, observe la figura 3:

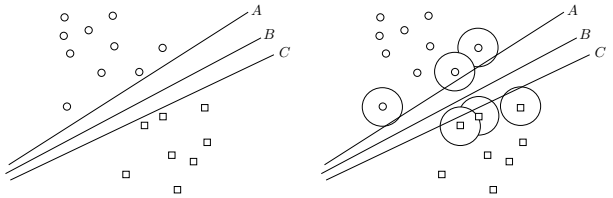

Figura 3: Los clasificadores A, B, C tienen error cero con respecto a los centros de los datos, pero sólo el clasificador 'B' tiene error cero si se tiene en cuenta la imprecisión de las mediciones.

Tres superficies de decisión A, B y C separan las dos categorías de la muestra, sin errores. Ahora bien, si cada dato se envuelve en un círculo, sólo la superficie B separa las dos clases, porque es la que más dista de los puntos más cercanos de la muestra. Esto parece indicar que la adición artificial de imprecisión a datos nítidos, y la optimización de un clasificador sobre los datos extendidos, podría producir una mejora en la capacidad de generalización. También podría existir una relación entre el diseño de un clasificador de margen mínimo y el diseño de un clasificador con datos imprecisos.

Por otra parte, en la figura 4 se muestra un problema de regresión, donde también podrían encontrarse ventajas. Tanto en la parte izquierda de la figura, como en la derecha, los modelos de regresión tienen error cuadrático medio cero. Sin embargo, si cada uno de los datos se envuelve en un intervalo, el error cuadrático es más específico en el modelo de la izquierda. De nuevo, la extensión de los datos sería equivalente a añadir al error cuadrático medio un término de regularización, lo que tal vez, como en el caso anterior, pueda explicarse estudiando la relación con las máquinas de vectores de soporte.

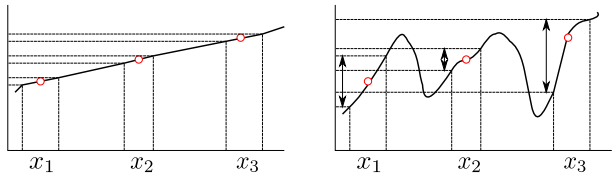

Figura 4: Ambos modelos de regresión tienen error cuadrático cero sobre los datos nítidos, pero el modelo de la izquierda tiene un error más específico sobre los datos extendidos, porque su pendiente en los puntos de la muestra es menor.

2.3. Problemática

El tratamiento de datos imprecisos es interesante tanto en problemas nítidos como en problemas con datos imprecisos, u otros que puedan codificarse mediante una pertenencia borrosa. Este tratamiento depende del concepto de dato vago (entendido como instancia de una VAB), y afecta al diseño del motor de inferencia del SBRD. Pero además, todas las fases del desarrollo del sistema necesitan adaptarse al nuevo tipo de datos: para diseñar un SBRD se necesitan nuevas técnicas de preprocesamiento y mejora de calidad de los datos, aprendizaje de bases de conocimiento y validación de resultados. Estos cambios se estudian en las siguientes secciones.

3. PREPROCESAMIENTO DE DATOS VAGOS

El preprocesado de datos imprecisos ha sido estudiado de forma desigual. Por una parte, no tenemos constancia de que se hayan desarrollado algoritmos de selección de instancias vagas. Por otra, hay varias extensiones borrosas de la selección de características y transformaciones de datos en problemas de clasificación y regresión. Ambas partes se discuten a continuación.

3.1. Selección de instancias

En problemas con datos precisos, la selección de instancias se basa en detectar las zonas de mayor densidad entre los ejemplos: una instancia puede ser redundante si existen en el fichero de datos otros ejemplos similares. Por el contrario, en problemas con datos vagos, es posible que una instancia intrínsicamente aporte poca información (como caso extremo, un ejemplo en que todos los datos de entrada son desconocidos siempre es redundante). Sería posible investigar la relación entre los diferentes costes de obtención de una instancia con distintas tolerancias, y el aporte de información que en cada caso esa instancia le haría a la base de conocimiento final.

3.2. Selección de características y transformaciones

En problemas de clasificación, puede determinarse el subconjunto más relevante de características mediante algoritmos basados en medidas de dependencia entre variables (como por ejemplo [27, 31]) o con técnicas de búsqueda (como [33]). En problemas de regresión es habitual emplear transformaciones, como el análisis de componentes principales y el análisis factorial (desarrollados en [15, 21] para datos borrosos). Las versiones borrosas de otros enfoques modernos, como el análisis de componentes independientes, están dirigidas a mejorar la robustez cuando se trabaja con datos nítidos [19] y no admiten datos vagos; su extensión es una tarea pendiente. Por el contrario, el escalado multidimensional sí ha sido generalizado al caso difuso [12, 16, 34], de forma que la proyección de los datos no sólo incluye información acerca de las similaridades entre las instancias sino también acerca de la vaguedad de cada una (ver figura 5).

4. DISEÑO DEL MOTOR DE INFERENCIA

La regla de inferencia composicional (extensión cilíndrica del dato de entrada, intersección con el grafo borroso codificado en la base de conocimiento, proyección en la salida) no necesariamente es compatible con la interpretación posibilística de un conjunto borroso vista en la sección 2. Para clarificar nuestra opinión, observe la siguiente base de conocimiento:

```
SI X ES PEQUEÑO ENTONCES Y ES GRANDE
SI X ES GRANDE ENTONCES Y ES PEQUEÑO
```

donde las pertenencias de los valores lingüísticos son $\texttt{PEQUEÑO}(t) = 1 - t$, y $\texttt{GRANDE}(t) = t$ para $t \in [0, 1]$. El grafo borroso de este sistema es

$$G(x, y) = \text{máx}\{\text{mín}\{x, 1 - y\}, \text{mín}\{1 - x, y\}\}. \quad (1)$$

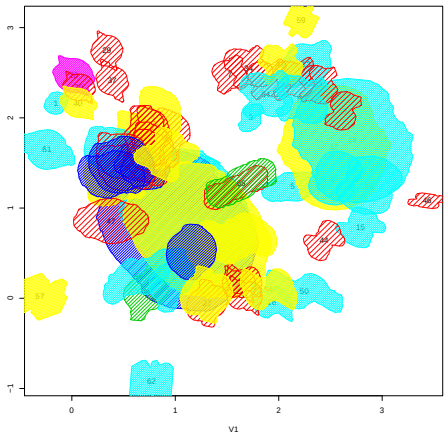

Figura 5: Extensión de MDS al caso borroso: se muestra gráficamente la similaridad entre las instancias y la vaguedad de cada una.

Por tanto, dada una entrada imprecisa X, la salida es el conjunto borroso

$$H_X(y) = \sup_x \text{mín}\{G(x, y), X(x)\}, \quad (2)$$

que no es normal (ver figura 6) y que por tanto no se corresponde con una distribución de posibilidad.

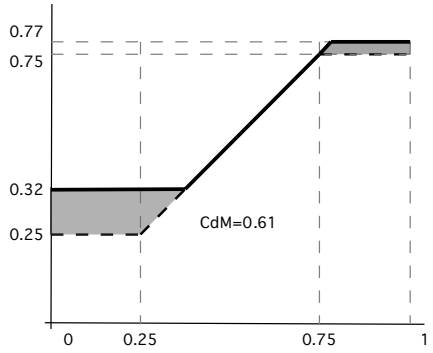

Figura 6: Inferencia con datos vagos (I): la curva con guiones es la imagen de la entrada crisp 0.25 mediante la regla de inferencia composicional. El valor defuzzificado es 0.62. La curva continua es la imagen del conjunto borroso triangular (0.15;0.25;0.35). El valor defuzzificado es 0.61

Para obtener una salida compatible con el modelo de imprecisión, puede extenderse la salida defuzzificada mediante el principio de extensión. Esta salida defuzzificada es

$$g(x_0) = \text{CdM}(H_{x_0}(y)) \quad (3)$$

donde CdM significa "centro de masas". La extensión tiene la forma que sigue:

$$F(X, y) = \sup\{\alpha \mid g^{-1}(y) \in X_\alpha\}. \quad (4)$$

La imagen de la misma entrada usada en la figura 6, ahora mediante la ecuación 4, puede verse en la figura 7. En esta segunda figura se aprecia que, con este tipo de inferencia, el que la entrada sea el valor vago $0{,}25 \pm 0{,}1$ implica que la salida tenga una tolerancia $\pm 0{,}05$.

El inconveniente de este método está, sin embargo, en la carga computacional que supone calcular la respuesta del sistema. Sería interesante desarrollar una versión de la regla de inferencia composicional, y tal vez un nuevo conjunto de operadores, que permitiesen producir una aproximación razonable al resultado mostrado en esta última figura. En este sentido, creemos que es posible que haya puntos en común entre los mecanismos de razonamiento de los SBRDs y los algoritmos utilizados para realizar inferencia en redes credales [11].

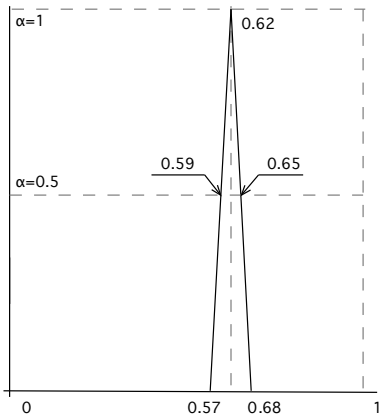

Figura 7: Inferencia con datos vagos (II): Imagen del conjunto borroso triangular (0.15;0.25;0.35) mediante la ecuación 4. El punto modal es la imagen del punto modal de la entrada.

5. PRECISION DE UN SBRD. OPTIMIZACIÓN GENÉTICA

La medición de la calidad de un SBRD se basa en la interpretabilidad lingüística y en la precisión del mismo. Sólo las medidas de precisión se ven afectadas cuando los datos son imprecisos. En ese caso, el modelo de VAB es determinante en la estimación de la precisión.

En general, se considerará que la precisión de un SBRD sobre datos vagos es, a su vez, un valor conocido de forma imprecisa. En trabajos anteriores [24, 28, 29] hemos adaptado la definición de la varianza de una VAB, que para el modelo de primer orden propuesto en la sección 2 se reduce a un intervalo [7, 10]. En [26] se propuso otro criterio específico para problemas de clasificación, que produce un número borroso discreto. En ambos casos, los algoritmos de aprendizaje de reglas borrosas que optimicen este criterio deben optimizar una función intervalo-valorada o borrosa [29, 30].

5.1. Sistemas borroso-genéticos

Los algoritmos genéticos han probado ser muy eficaces para solucionar el problema del aprendizaje de un SBRD a partir de datos nítidos [6, 17, 18]. En trabajos anteriores, hemos conseguido obtener algunos resultados que permiten formular sistemas borroso-genéticos a partir de información vaga, siguiendo los enfoques Pittsburgh [24], Iterativo [25] y GCCL [32] en problemas de modelado.

Por el contrario, hasta el momento no se ha publicado ninguna implementación del cuarto enfoque, los algoritmos tipo Michigan. Tampoco tenemos conocimiento de que los sistemas borroso-genéticos hayan sido aplicados a problemas de clasificación con datos imprecisos.

6. VALIDACION

El uso de medidas de calidad intervalares o borrosas plantea algunos problemas con la comparación estadística de los resultados de diferentes algoritmos. Aunque existen numerosos trabajos teóricos de estadística borrosa acerca de test estadísticos según el modelo clásico de probabilidad, los estudios relacionados con los modelos de primer o segundo orden son menos numerosos [13, 20, 35]. Es interesante observar que estos dos últimos modelos, cuando la observaciones son intervalo-valoradas, guardan relación con los trabajos en probabilidades imprecisas [3] y con algunos enfoques recientes en modelos econométricos [1]. Recientemente, se está trabajado en la base teórica de los tests de remuestreo que se podrían aplicar en un futuro a la comparación de las calidades de diferentes algoritmos [8, 9].

Por otro lado, es interesante destacar que no existe un catálogo de problemas sobre los que se puedan evaluar las propiedades de diferentes algoritmos. Hasta la fecha, cada autor propone sus propios problemas de validación. Sería deseable que se elaborase un catálogo de problemas con diferentes complejidades y con resultados conocidos.

7. CONCLUSIONES Y PROBLEMAS ABIERTOS

En este estudio hemos mostrado las líneas generales del uso de datos vagos con sistemas basados en reglas borrosas. Se ha propuesto combinar las herramientas de la lógica borrosa con las de la estadística borrosa, aprovechando resultados recientes acerca de la representación de la imprecisión en la observación de los

resultados de un experimento aleatorio.

La principal conclusión de este estudio es que muchos de los problemas del uso de datos vagos con SBRDs están sin resolver de forma satisfactoria. A modo de resumen, los retos mencionados en este trabajo han sido:

1. El estudio teórico de las relaciones entre los problemas con imprecisión añadida y las máquinas de vectores de soporte.
2. El desarrollo de las técnicas de análisis de componentes principales e independientes para datos imprecisos.
3. Nuevos algoritmos de selección de características e instancias para datos vagos.
4. Nuevos mecanismos de razonamiento borroso que permitan aproximar la solución basada en el principio de extensión.
5. Otras medidas de calidad para clasificadores y modelos.
6. Estudio teórico y práctico de algoritmos de optimización de funciones intervalo-valoradas y borrosas.
7. Tests estadísticos que permitan comparar la calidad de dos algoritmos sobre datos vagos.
8. Bancos de prueba para comparar los nuevos algoritmos.

Agradecimientos

Este trabajo ha sido soportado por el Ministerio de Ciencia e Innovación, a través del proyecto de investigacón TIN2005-08036-C05-05.

Referencias

[1] Beresteanu, A., and Molinari, F.: Asymptotic Properties for a Class of Partially Identified Models, Econometrica, 2008 (En prensa)

[2] Baudrit, C., Couso, I., Dubois, D. Joint propagation of probability and possibility in risk analysis: Towards a formal framework, International Journal of Approximate Reasoning, 45, pp 82–105. 2007.

[3] Baudrit, C., Dubois, D., Perror, N. Representing parametric probabilistic models tainted with imprecision. Fuzzy Sets and Systems, 2008 (En prensa)

[4] Casillas, J, Cordon, O., Herrera, F., Magdalena, L. Interpretability issues in fuzzy modeling. Springer. Heidelberg, Germany. 2003.

[5] Castro, J., Delgado, M. Fuzzy Systems with Defuzzification are Universal Approximators. IEEE Trans. on Syst., Man, and Cybern 26(1), pp 149–152. 1996.

[6] Cordón, O., Herrera, F., Hoffmann, F., Magdalena, L. Genetic fuzzy systems: Evolutionary tuning and learning of fuzzy knowledge bases. World Scientific Publishing Company, Singapore. 2001.

[7] Couso, I., Dubois, D. Montes, S., Sánchez, L.,On various definitions of the variance of a fuzzy random variable. ISIPTA 07 - Fifth International Symposium on Imprecise Probability: Theories and Applications, 135-144. 2007

[8] Couso, I., Sánchez, L., Mark-recapture techniques in statistical tests for imprecise data. International Journal of Approximate Reasoning (sometido)

[9] Couso, I., Sánchez, L., Defuzzification of fuzzy p-values. Fourth International Workshop on Soft methods in probability and statistics SPMS 08 (sometido)

[10] Couso, I., Sánchez, L. Higher Order Models for Fuzzy Random Variables. Fuzzy Sets and Systems 159 (3) pp. 237-258. 2008

[11] Ide, J. S., Cozman, F. G., Approximate algorithms for credal networks with binary variables. International Journal of Approximate Reasoning 48(1), pp 275-296. 2008.

[12] Denoeux, T., Masson, M. H., Multidimensional scaling of interval-valued dissimilarity data. Pattern Recognition Letters 21, pp 83-92. 2000.

[13] Denoeux, T, Masson, M. H. , Hebert, P.A. Nonparametric Rank-Based Statistics and Significance Test for Fuzzy Data, Fuzzy Sets and Systems, 153. pp 1-28. 2005.

[14] Ferson, S., Kreinovich, V., Hajagos, J., Oberkampf, W., Ginzburg, L., Experimental Uncertainty Estimation and Statistics for Data Having Interval Uncertainty. Tech. Rep. SAND2007-0939, Sandia National Laboratories, 2007.

[15] Giordani, P., Kiers, H., Principal Component Analysis of symmetric fuzzy data. Computational Statistics and Data Analysis 45(3). pp 519-548. 2004.

[16] Hebert, P. A., Masson, M. H., Denoeux, T. Fuzzy multidimensional scaling. Computational Statistics and Data Analysis 51. pp 335-359. 2006.

[17] Herrera, F. Genetic Fuzzy Systems: Status, Critical Considerations and Future Directions. International Journal of Computational Intelligence Research. 1 (1). pp 59-67. 2005.

[18] Herrera, F. Genetic Fuzzy Systems: Taxonomy, Current Research Trends and Prospects. Evolutionary Intelligence 1. pp 27-46. 2008.

[19] Honda, K., Ichihashi, H. Fuzzy local independent component analysis with external criteria and its application to knowledge discovery in databases. International Journal of Approximate Reasoning 42(3), pp. 159-173. 2006.

[20] Kruse, R., Meyer, K. D., Statistics with vague data. D. Reidel Publishing Company. 1987.

[21] Nakamori, Y., Watada, J. Factor analysis for fuzzy data. Sixth Intl. Conf. Fuzzy Systems, pp. 1115-1120. 1997.

[22] Otero, A., Otero, J., Sanchez, L., Villar, J. R. Longest path estimation from inherently fuzzy data acquired with GPS using genetic Algorithms. Proc. of the 2006 IEEE International Conference on Evolutionary Fuzzy Systems, Ambleside, UK. pp 30-305. 2006.

[23] Puri, M. L., Ralescu, D. Fuzzy random variables. J. Math. Anal. Appl., 114 pp 409-422, 1986.

[24] Sánchez, L. Couso, I. Casillas, J. A multiobjective genetic fuzzy system with imprecise probability fitness for vague data. Proc. of the 2nd International Symposium on Evolving Fuzzy Systems (EFS 2006), 131-136, 2006

[25] Sánchez, L., Otero, J., Villar, J. R. Boosting of fuzzy models for high-dimensional imprecise datasets. Proc. of the 2006 International Conference in Information Processing and Management of Uncertainty in Knowledge-based Systems, Paris, France. pp 1965-1973. 2006.

[26] Sánchez, L., Couso, I., Advocating the use of imprecisely observed data in Genetic Fuzzy Systems. IEEE Trans Fuzzy Sys 15(4). pp 551-562. 2007.

[27] Sánchez, L.; Suarez, M. R.; Villar, J. R.; Couso, I. Some results about mutual information-based feature selection and fuzzy discretization of vague data. FUZZ-IEEE 2007. pp 23-26. 2007

[28] Sánchez, L., Otero, J., Learning Fuzzy Linguistic Models from low quality data by genetic algorithms. FUZZ-IEEE 2007. pp 23-26. 2007

[29] Sánchez, L. Couso, I. Casillas, J. Modelling Vague Data with Genetic Fuzzy Systems under a Combination of Crisp and Imprecise Criteria, in Proc. of the 2007 IEEE Symposium on Computational Intelligence in Multicriteria Decision Making (MCDM'2007), pp. 30–37, Honolulu, Hawaii, USA. 2007

[30] Sánchez, L., Villar, J. R. Obtaining transparent models of chaotic systems with multiobjective simulated annealing algorithms. Information Sciences 178(4). pp 952-970. 2008

[31] Sánchez, L., Suárez, M. R., Villar, J. R., Couso, I. Mutual Information-based Feature Selection and Fuzzy Discretization of Vague Data. International Journal of Approximate Reasoning (Admitido)

[32] Sánchez, L., Otero, J., Couso, I. Obtaining Linguistic Fuzzy Rule-based Regression Models from Imprecise Data with Multiobjective Genetic Algorithms. Soft Computing (Sometido)

[33] Sánchez, L., Villar, J. R., Couso, I. Genetic Feature Selection for Low Quality Data. IPMU08 (Admitido). Torremolinos, Málaga. 2008

[34] Sánchez, L., Palacios, A., Suárez, A., Couso, I. Selecting the most relevant tests from vague data in the early diagnosis of dyslexia IPMU08 (Admitido). Torremolinos, Málaga. 2008

[35] Zaffalon, M. Conservative rules for predictive inference with incomplete data, in: Fourth International Symposium on Imprecise Probabilities and their Applications, pp. 406-415. 2005.