

Análisis descriptivo mediante aprendizaje supervisado basado en patrones emergentes

C.J. Carmona¹, F. Pulgar², A.M. García², P. González², M.J. del Jesus²

Área de Lenguajes y Sistemas del Departamento de Ingeniería Civil, Universidad de Burgos ¹

`cjcarmona@ubu.es`

Departamento de Informática, Universidad de Jaén ²

Resumen La búsqueda patrones emergentes persigue describir un problema mediante la obtención de tendencias emergentes en el tiempo, o la caracterización de diferencias entre clases o entre un grupo de variables. En este trabajo se presenta un nuevo modelo evolutivo para la obtención de patrones emergentes basado en un enfoque mono-objetivo, con la capacidad de extraer patrones emergentes para todos los valores de la variable objetivo. Los resultados del estudio experimental muestran el poder descriptivo del nuevo modelo frente al algoritmo clásico de minería de patrones emergentes. Los resultados también muestran la gran capacidad de esta técnica para describir conocimiento en problemas basados en aprendizaje supervisado.

Keywords: Patrones emergentes, sistemas difusos evolutivos, minería de datos descriptiva, aprendizaje supervisado

1. Introducción

En el proceso de minería de datos se puede distinguir entre inducción predictiva cuyo objetivo es descubrir conocimiento para clasificar o predecir [15], e inducción descriptiva cuyo principal objetivo es extraer conocimiento interesante de los datos sin la necesidad de trabajar con datos etiquetados [1]. Sin embargo, en los últimos años ha habido un gran interés por parte de la comunidad investigadora en el estudio de modelos descriptivos mediante aprendizaje supervisado.

En minería de patrones emergentes, es muy importante el concepto de soporte, definido como el porcentaje de ejemplos del conjunto de datos que satisfacen la regla o patrón. De esta forma, la minería de patrones emergentes tiene como objetivo buscar patrones cuyo soporte varíe mucho de un conjunto de datos a otro [5]. Por definición, este concepto se puede extrapolar para medir la diferencia de soporte en un mismo problema pero respecto a los distintos valores de la variable objetivo, es decir, medir diferencias de soporte con respecto a los distintos valores de una variable de interés. Este concepto ha recibido especial interés en el campo de la inducción predictiva, pero no así dentro de la inducción descriptiva. Su gran carácter diferenciador y la posibilidad de utilizarlo como

método descriptivo y/o exploratorio, nos ha llevado a presentar en este trabajo un primer estudio sobre su eficacia, tal y como se adelantaba en [13].

En concreto, en esta contribución se presenta un nuevo enfoque evolutivo con la capacidad de obtener patrones emergentes para todos los valores de la variable objetivo. El enfoque se denomina EvAEP y es un algoritmo basado en un sistema evolutivo difuso [9]. Estos sistemas están basados en algoritmos evolutivos [7] que ofrecen un amplio conjunto de ventajas para la extracción de conocimiento y en concreto para procesos de inducción de reglas. Además, emplean la lógica difusa [18], que considera la utilización de conjuntos difusos con etiquetas lingüísticas para representar la información permitiendo obtener un conocimiento muy cercano al razonamiento humano. También, permiten mejorar la interpretabilidad de los modelos obtenidos para la búsqueda de patrones emergentes, y en general en el análisis de datos para establecer relaciones e identificar patrones [12].

Este trabajo se organiza en las siguientes secciones: En la Sección 2, se analiza el contexto del problema y en concreto la definición, características y técnicas de los patrones emergentes presentes a lo largo de la literatura, en la Sección 3 se presenta un nuevo enfoque difuso evolutivo para la obtención de patrones emergentes y en la Sección 4 se presenta el estudio experimental para la comparación del enfoque de minería de patrones emergentes clásico y el nuevo enfoque presentado en este trabajo; por último, se presentan las conclusiones obtenidas en el trabajo junto con los trabajos futuros.

2. Minería de patrones emergentes

La minería de datos se centra en desarrollar y aplicar algoritmos que, bajo limitaciones aceptables de eficiencia computacional, permitan la obtención de patrones sobre los datos. Esta obtención de patrones puede abordarse mediante dos objetivos básicos:

- **Predicción.** El sistema busca patrones para predecir un comportamiento futuro. Los modelos desarrollados dentro de este objetivo están basados en aprendizaje supervisado. Entre las tareas predictivas se pueden encontrar por ejemplo: clasificación, regresión o análisis de series temporales.
- **Descripción.** El sistema busca patrones para presentarlos a un experto en una forma comprensible para él, y que describen y aportan información de interés sobre el problema y el modelo que subyace bajo los datos. Los modelos desarrollados en este enfoque se basan en aprendizaje no supervisado. Entre las tareas descriptivas se pueden encontrar por ejemplo: agrupamiento, sumarización o asociación.

A pesar de existir una clara distinción entre las técnicas dependiendo de los objetivos que pretenden abordar, en la actualidad existen un conjunto de técnicas que se encuentran a medio camino entre las técnicas predictivas y las descriptivas, agrupadas bajo el nombre de descubrimiento de reglas descriptivas basadas en aprendizaje supervisado [13], que intentan obtener reglas o conjuntos

de ítems de una categoría o clase prefijada para describir información significativa y relevante del conjunto de datos.

Su principal objetivo no es clasificar nuevas instancias, sino comprender o encontrar fenómenos subyacentes; es decir, encontrar información desconocida u oculta difícil de descubrir por los expertos, sobre el valor de una clase prefijada o variable de interés. Dentro de este grupo de técnicas se encuentran la: minería de conjuntos de contraste [3], el descubrimiento de subgrupos [10] y los patrones emergentes [6].

La minería de patrones emergentes busca conocimiento relacionado con los distintos valores de la variable objetivo, donde el número de instancias cubiertas por un patrón sea muy elevado para un valor de la variable objetivo y muy bajo (o cero) para el resto; es decir, que el mismo patrón tenga un soporte muy alto para una clase y muy bajo para las demás.

Los patrones emergentes fueron formulados inicialmente por Dong y Li [5,6], de la siguiente forma:

Dado un patrón x cualquiera se denomina como emergente si el índice de crecimiento del mismo entre dos conjuntos de datos (D_1 y D_2) es superior a uno.

El índice de crecimiento viene definido de la siguiente forma:

$$IndiceCrecimiento(x) = \begin{cases} 0, & Si \ Sop_1(x) = Sop_2(x) = 0, \\ \infty, & Si \ Sop_1(x) = 0 \wedge Sop_2(x) \neq 0, \\ \frac{Sop_2(x)}{Sop_1(x)}, & en \ otro \ caso \end{cases} \quad (1)$$

siendo el $Sop_1(x)$ el soporte del patrón x en el conjunto de datos D_1 y $Sop_2(x)$ el soporte del mismo patrón en el conjunto de datos D_2 , o lo que es lo mismo, el soporte de un patrón para un valor de la variable objetivo con respecto a otro valor de la variable objetivo.

Debido a esta propiedad, esta técnica es capaz de diferenciar o clasificar instancias para una clase u otra mediante el índice de crecimiento; de hecho un modelo obtenido por un algoritmo de este tipo puede ser utilizado tanto con carácter descriptivo, para describir ciertos comportamientos en los datos, como con carácter predictivo que permite predecir nuevas instancias que se deseen incorporar en el modelo obtenido.

La representación del conocimiento en la minería de patrones emergentes puede ser por tanto realizada mediante reglas, de la siguiente forma:

$$Cond \longrightarrow Target_{value} \quad (2)$$

donde $Cond$ representa una conjunción de pares atributo-valor y $Target_{value}$ representa un valor para la variable objetivo, es decir, todas las reglas obtenidas para ese valor de la variable objetivo miden el índice de crecimiento frente a otro/s valores de la variable objetivo. Este tipo de representación permite ofrecer a los expertos una mayor interpretabilidad de los resultados puesto que las

reglas aportan información directa sobre el problema, de forma comprensible sin necesidad de una posterior transformación.

La mayoría de trabajos presentes en la literatura han utilizado la minería de patrones emergentes como un mero clasificador, pero no han explotado las posibilidades que esta técnica tiene de cara a describir información y conocimiento relacionado con la ganancia de información que un patrón puede tener entre distintos valores de la variable objetivo. Entre los distintos algoritmos que se han presentado en la literatura hay que destacar el algoritmo DeEPs [14]. Este algoritmo realiza una búsqueda de patrones emergentes basados en el concepto de límites [5]. Para cada instancia se obtienen todos aquellos patrones emergentes que están dentro de los límites y con estos patrones se puntúa al valor de la clase a la que pertenece. Ya que el número de patrones es excesivamente alto, sobre todo en bases de datos con muchas instancias, es necesario realizar un filtrado de patrones emergentes por longitud o por frecuencia.

Para medir la calidad de los patrones emergentes obtenidos por un algoritmo, y en definitiva, la calidad de un algoritmo de minería de patrones emergentes, es necesario analizar un amplio abanico de medidas de calidad utilizadas a lo largo de la literatura [8]. Sin embargo, para simplificar el análisis nos centraremos en los siguientes objetivos fundamentales:

- Simplicidad, que se mide mediante el número de patrones y el número de variables que conforman el patrón.
- Poder discriminativo, que se puede medir mediante el índice de crecimiento y la confianza de los patrones.
- Generalidad, que se mide mediante la sensibilidad y el soporte.
- Ganancia de información, que se puede medir mediante la atipicidad y la ganancia del patrón.

3. EvAEP: Evolutionary Algorithm for extracting fuzzy Emerging Patterns

Esta sección describe un nuevo modelo difuso evolutivo, EvAEP, para extraer patrones emergentes difusos que describan problemas desde un nuevo punto de vista bajo el aprendizaje supervisado.

El principal objetivo de EvAEP es la extracción un número variable de reglas o patrones que describan información con respecto a una variable objetivo. Considerando que una variable objetivo podrá tener distintos valores, el algoritmo es capaz de obtener patrones para todos los valores de esa variable objetivo al ejecutarse una vez para cada valor. Dentro del aprendizaje automático existen un gran número de problemas a resolver que normalmente se caracterizan por tener un diferente número de valores en la variable objetivo; es decir, podemos tener problemas con dos o más valores de la variable objetivo. Para realizar el análisis descriptivo de los patrones emergentes en un problema se pueden tener en cuenta dos metodologías: Uno contra uno o uno contra todos. Esta última es la metodología utilizada por este algoritmo.

El algoritmo evolutivo emplea una codificación de ‘Cromosoma = Regla’ donde sólo se representa la parte del antecedente de la regla. El antecedente está compuesto por una conjunción de pares atributo-valor, indicando mediante un valor especial aquellas variables que no forman parte de la regla; es decir, un cromosoma representa todas las variables de entrada del problema. La Fig. 1 muestra la representación de una regla para el algoritmo EvAEP.

$$\begin{array}{c} \textit{Genotipo} \\ \left| \begin{array}{c|c|c|c} x_1 & x_2 & x_3 & x_4 \\ \hline 3 & \emptyset & 1 & \emptyset \end{array} \right| \Rightarrow \text{ IF } (x_1 = 3) \text{ Y } (x_3 = 1) \text{ THEN } (x_{Obj} = \textit{ValorObjetivo}) \\ \textit{Fenotipo} \end{array}$$

Figura 1. Representación de una regla en EvAEP

Cuando las variables tienen rangos continuos, el modelo emplea reglas difusas y los conjuntos difusos correspondientes a las etiquetas lingüísticas se definen por medio de funciones de pertenencia triangulares uniformes.

El algoritmo utiliza un enfoque mono-objetivo de aprendizaje de reglas iterativo [17] que es ejecutado una vez por cada valor de la variable objetivo, es decir, para cada valor de la variable objetivo se obtienen de forma iterativa los mejores individuos o patrones emergentes del problema. El algoritmo itera para obtener patrones emergentes hasta que encuentra un patrón que no es emergente, o encuentra un patrón que no cubre nuevos ejemplos del problema, o, todos los ejemplos para ese valor de la variable objetivo ya se encuentran cubiertos por los patrones extraídos.

El esquema de funcionamiento del algoritmo EvAEP se muestra en la Fig. 2.

El objetivo propiedad del algoritmo es la obtención de patrones emergentes precisos que tengan la capacidad de aportar una alta generalidad y ganancia de información. A continuación se describen los elementos principales del algoritmo.

3.1. Inicialización sesgada

EvAEP genera una población inicial (P_0) cuyo tamaño se determina mediante un parámetro. El objetivo de esta función de inicialización es crear de una parte de los individuos de forma que utilicen como máximo un porcentaje de las variables. En concreto, el algoritmo crea una población con un 50% de individuos generados de forma completamente aleatoria y el resto de la población con individuos que tienen inicializada como mínimo una variable y como máximo un 80% de las variables de entrada del problema.

Este operador permite obtener una primera población en la que parte de los individuos utilizan un bajo porcentaje de variables, lo que aporta generalidad a los patrones obtenidos.

3.2. Operadores genéticos

La población de la siguiente generación es generada mediante distintos operadores genéticos ampliamente utilizados en la literatura.

```

INICIO
Conjunto de patrones emergentes =  $\emptyset$ 
repeat
  repeat
    Generar P(0)
    Evaluar P(0)
  repeat
    Incluir el mejor individuo en P(nGen+1)
    Completar P(nGen+1): Cruce y Mutación de individuos de P(nGen)
    Evaluar P(nGen+1)
    nGen  $\leftarrow$  nGen + 1
  until Número de evaluaciones alcanzado
  Obtener la mejor regla (R)
  Patrones emergentes  $\cup$  R
  Marcar los ejemplos cubiertos por R
  until ÍndiceCrecimiento(R)  $\leq$  1 and R no cubre nuevos ejemplos para el objetivo
until Valores objetivo =  $\emptyset$ 
return Patrones emergentes
FIN

```

Figura 2. Esquema de funcionamiento de EvAEP

El algoritmo emplea elitismo de tamaño uno, de forma que el mejor individuo se almacena directamente en la población de la siguiente generación. El mejor individuo es aquél que obtenga el máximo valor en la siguiente función de agregación, y en caso de empate, el individuo con menor número de variables:

$$NSup(R) * 0,5 + Fitness(R) * 0,5 \quad (3)$$

donde $NSup(R)$ equivale al número de ejemplos cubiertos por la regla R no cubiertos hasta el momento por los anteriores patrones, entre el número de ejemplos que quedan por cubrir de la variable objetivo. La función $Fitness$ es la función de evaluación de los individuos, que se detalla en la siguiente sección.

Por otro lado, el algoritmo emplea el operador de cruce multi-punto [11] y un operador de mutación sesgada que fue introducido en un algoritmo de descubrimiento de subgrupos [4] para crear nuevos individuos en la nueva población. La selección de los individuos que se cruzarán y mutarán se realiza mediante una selección por torneo [16].

3.3. Función de evaluación

Es uno de los aspectos fundamentales del algoritmo y busca obtener patrones emergentes que tengan una alta confianza o precisión, con la máxima generalidad posible y una ganancia de información relevante. El algoritmo EvAEP utiliza la siguiente medida de fitness para los individuos:

$$Fitness(R) = \sqrt{TPr * TNr} \quad (4)$$

que se corresponde con la media geométrica del individuo que intentan maximizar la precisión en las dos clases que componen un determinado problema de la forma más balanceada posible. El primer componente (*TPr*) es conocido como *True Positive rate* o sensibilidad y es el porcentaje de patrones correctamente clasificados de la clase positiva con respecto al número total de ejemplos del valor de la variable objetivo. Por otro lado, el segundo componente (*TNr*) conocido como *True Negative rate* es la capacidad o efectividad para reconocer los elementos de los valores de la variable objetivo que no corresponden al valor que estamos analizando.

4. Estudio experimental

En esta contribución se presenta un estudio preliminar de la capacidad descriptiva de los patrones emergentes y de la calidad de la propuesta presentada con respecto al algoritmo clásico por excelencia de los patrones emergentes, DeEPs. En concreto se realiza un estudio con distintos conjuntos de datos del repositorio KEEL [2]: Australian, Balance, Breast, Glass, Hypo, Iris, Pima, Votes y Wine. Es importante destacar que entre estos conjuntos se encuentran algunos con dos valores para la variable objetivo y otros con más de dos valores, y con conjuntos con variables discretas y/o continuas.

La comparación entre los algoritmos se ha realizado mediante una validación cruzada de 10 particiones y debido al carácter no determinista del algoritmo evolutivo se realizan 3 ejecuciones por cada par entrenamiento-test del conjunto de datos. De esta forma, la media de estas ejecuciones es la que se obtiene como resultado, es decir; los resultados muestran los promedios de las medias obtenidas en cada una de las ejecuciones por los patrones. Los parámetros empleados por los algoritmos se pueden observar en la Tabla 1.

Tabla 1. Parámetros de los algoritmos empleados en el estudio experimental

Algoritmo	Parámetros
DeEPs	$\alpha=0.12$, número máximo de patrones=3, filtrado=longitud
EvAEP	Número de etiquetas lingüísticas=3, máximo número de evaluaciones=10000, tamaño de la población=100, porcentaje de cruce=60 %, porcentaje de mutación=1 %

La Tabla 2 detalla los promedios de los patrones obtenidos por los algoritmos (*Alg.*), en los distintos conjuntos de datos analizados (*BD*). Las medidas de calidad analizadas son: el número de patrones o reglas (*Patrones*), el número de variables en el antecedente (*Vars*), la longitud (*Long*), *Ganancia*, Sensibilidad (*Sens*) y la Confianza (*Conf*). Hay que considerar que los patrones totales obtenidos por DeEPs se indican entre paréntesis. Todas las medidas de calidad analizadas pueden ser consultadas en [8].

Tabla 2. Resultados obtenidos en el estudio experimental

<i>BD</i>	<i>Alg</i>	<i>Patrones</i>	<i>Vars</i>	<i>Long</i>	<i>Ganancia</i>	<i>Sens</i>	<i>Conf</i>
Australian	EvAEP	6.10	2.66	0.064	0.294	0.682	0.736
	DeEPs	6.00 (6655)	4.60	0.534	0.031	0.056	0.803
Balance	EvAEP	9.70	1.44	0.108	0.204	0.388	0.614
	DeEPs	7.10 (1026)	2.67	0.165	0.014	0.026	0.185
Breast	EvAEP	17.80	1.12	0.222	0.189	0.323	0.784
	DeEPs	6.00 (3396)	4.96	0.148	0.239	0.399	1.000
Glass	EvAEP	9.70	4.98	0.256	0.360	0.267	0.328
	DeEPs	17.70 (236)	6.13	0.234	0.141	0.077	0.195
Hypo	EvAEP	4.80	5.02	0.120	0.083	0.652	0.877
	DeEPs	6.00 (3671)	4.98	0.208	0.121	0.337	0.791
Iris	EvAEP	3.20	1.60	0.312	0.711	0.775	0.797
	DeEPs	9.00 (123)	3.40	0.490	0.361	0.380	0.835
Pima	EvAEP	7.90	2.73	0.137	0.142	0.438	0.574
	DeEPs	6.00 (1125)	4.28	0.332	0.033	0.077	0.739
Votes	EvAEP	5.50	1.42	0.054	0.409	0.766	0.815
	DeEPs	6.00 (7749)	8.01	0.249	0.130	0.232	0.924
Wine	EvAEP	7.20	1.90	0.185	0.666	0.783	0.607
	DeEPs	9.00 (3774)	5.72	0.267	0.042	0.042	0.233
<i>AVERAGE</i>	EvAEP	7.98	2.54	0.169	0.339	0.563	0.681
	DeEPs	8.08 (3083)	4.97	0.291	0.123	0.180	0.633

La cantidad y variedad de medidas de calidad utilizadas en la comparación hace difícil obtener conclusiones generales de los resultados de la experimentación. Por ello, se analizan a continuación de forma separada los resultados para las distintas medidas de calidad presentadas en este estudio:

- El *número de patrones* obtenido por el algoritmo EvAEP es muy inferior al obtenido por el algoritmo clásico DeEPs. Sin embargo, con la utilización del filtrado de aquellos patrones con mayor sensibilidad para cada valor de la variable objetivo, se consigue reducir mucho el número de patrones emergentes obtenidos por DeEPs. A pesar de ello, el algoritmo evolutivo sigue obteniendo menor número promedio de patrones, lo que facilita su interpretación por parte de los expertos.
- El *número de variables* es bastante inferior en el algoritmo evolutivo con respecto al clásico. En este sentido, el algoritmo presentado en este trabajo obtiene patrones emergentes más generales e interpretables.
- La *longitud* mide la inversa del número de ejemplos cubiertos, por lo que un mejor resultado vendrá dado por aquellos algoritmos cuya longitud tienda a 0. En este sentido, EvAEP obtiene mejores resultados en promedio y en todos los conjuntos de datos salvo *Breast* y *Glass*.
- La *ganancia* es una medida de calidad clave dentro de la búsqueda de patrones emergentes y en este estudio preliminar el algoritmo evolutivo demuestra un mejor comportamiento de cara a obtener patrones con una mayor ganancia de información.

- En el caso de la *sensibilidad* puede determinar la calidad de los patrones emergentes descubiertos, ya que mide el número de ejemplos cubiertos correctamente para el valor de la variable objetivo. En este estudio, el algoritmo EvAEP consigue obtener valores de sensibilidad mucho más altos que DeEPs con una media cercana al 60 %.
- Para finalizar, la *confianza* mide la precisión de los patrones y en este sentido a pesar de obtener valores muy parejos entre ambos algoritmos, EvAEP tiene un comportamiento más homogéneo con un mejor promedio. Además, EvAEP obtiene mejor relación entre sensibilidad-confianza que es clave para describir buenos patrones emergentes con gran generalidad.

5. Conclusiones y trabajos futuros

En esta contribución se han aportado dos nuevos conceptos dentro de la minería descriptiva basada en aprendizaje supervisado:

- Un nuevo modelo evolutivo para la obtención de patrones emergentes, que demuestra en el estudio experimental realizado el gran aporte con respecto a un modelo clásico. Además, EvAEP aporta patrones con mayor simplicidad, mayor generalidad-precisión y ganancia de información.
- La utilización de la minería de patrones emergentes como método exploratorio y/o descriptivo, capaz de obtener información relevante con respecto a una variable de interés en un problema.

Para trabajos futuros, se plantea la necesidad de profundizar en el estudio y análisis de los modelos de minería de patrones emergentes existentes en la literatura, la necesidad de realizar un estudio detallado sobre la utilización de las medidas de calidad, así como el desarrollo de nuevos modelos.

Acknowledgments.

Este trabajo ha sido subvencionado por el Ministerio de Economía y Competitividad bajo el proyecto TIN2012-33856, Fondos FEDER.

Referencias

1. R. Agrawal, T. Imieliski, and A. Swami, *Mining association rules between sets of items in large databases*, Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data, ACM Press, 1993, pp. 207–216.
2. J. Alcalá-Fdez, A. Fernández, J. Luengo, J. Derrac, S. García, L. Sánchez, and F. Herrera, *KEEL Data-Mining Software Tool: Data Set Repository, Integration of Algorithms and Experimental Analysis Framework*, Journal of Multiple-Valued Logic and Soft Computing **17** (2011), no. 2-3, 255–287.
3. S. Bay and M. Pazzani, *Detecting Group Differences: Mining Contrast Sets*, Data Mining and Knowledge Discovery **5** (2001), 213–246.

4. C. J. Carmona, P. González, M. J. del Jesus, and F. Herrera, *NMEEF-SD: Non-dominated Multi-objective Evolutionary algorithm for Extracting Fuzzy rules in Subgroup Discovery*, IEEE Transactions on Fuzzy Systems **18** (2010), no. 5, 958–970.
5. G. Z. Dong and J. Y. Li, *Efficient Mining of Emerging Patterns: Discovering Trends and Differences*, Proceedings of the 5th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, ACM Press, 1999, pp. 43–52.
6. ———, *Mining border descriptions of emerging patterns from dataset pairs*, Knowledge and Information Systems **8** (2005), no. 2, 178–202.
7. A. E. Eiben and J. E. Smith, *Introduction to evolutionary computation*, Springer, 2003.
8. M. García-Borroto, O. Loyola-González, J.F. Martínez, and J.A. Carrasco-Ochoa, *Comparing quality measures for contrast pattern classifiers.*, Lecture Notes in Computer Science, vol. 8258, Springer, 2013, pp. 311–318.
9. F. Herrera, *Genetic fuzzy systems: taxonomy, current research trends and prospects*, Evolutionary Intelligence **1** (2008), 27–46.
10. F. Herrera, C. J. Carmona, P. González, and M. J. del Jesus, *An overview on Subgroup Discovery: Foundations and Applications*, Knowledge and Information Systems **29** (2011), no. 3, 495–525.
11. J. H. Holland, *Adaptation in natural and artificial systems*, University of Michigan Press (1975).
12. E. Hüllermeier, *Fuzzy methods in machine learning and data mining: Status and prospects*, Fuzzy Sets and Systems **156** (2005), no. 3, 387–406.
13. P. Kralj-Novak, N. Lavrac, and G. I. Webb, *Supervised Descriptive Rule Discovery: A Unifying Survey of Contrast Set, Emerging Pattern and Subgroup Mining*, Journal of Machine Learning Research **10** (2009), 377–403.
14. J. Y. Li, G. Z. Dong, K. Ramamohanarao, and L. Wong, *DeEPs: A New Instance-Based Lazy Discovery and Classification System*, Machine Learning **54** (2004), no. 2, 99–124.
15. D. Michie, D. J. Spiegelhalter, and C. C. Tayloy, *Machine Learning*, Ellis Horwood, 1994.
16. B. L. Miller and D. E. Goldberg, *Genetic Algorithms, Tournament Selection, and the Effects of Noise*, Complex System **9** (1995), 193–212.
17. G. Venturini, *SIA: A Supervised Inductive Algorithm with Genetic Search for Learning Attributes based Concepts*, Proceedings European Conference on Machine Learning, LNAI, vol. 667, Springer, 1993, pp. 280–296.
18. L. A. Zadeh, *The concept of a linguistic variable and its applications to approximate reasoning. Parts I, II, III*, Information Science **8-9** (1975), 199–249,301–357,43–80.