

Grupo de Sistemas Inteligentes (UPV-EHU)

P. Larrañaga, J.A. Lozano, I. Inza, E. Bengoetxea, T. Miquelez, A. Mendiburu, J.L. Jimenez, T. Romero, C. Gonzalez, R. Blanco, J.L. Flores, G. Santafé, and A. Perez

Departamento de Ciencias de la Computación e Inteligencia Artificial
Universidad del País Vasco - Euskal Herriko Unibertsitatea
<http://www.sc.ehu.es/iscg>

El Grupo de Sistemas Inteligentes de la UPV-EHU se constituyó en el año 1996. Desde entonces su labor se ha centrado en el desarrollo de metodología en las áreas del aprendizaje automático y la computación evolutiva. Algunos de los métodos desarrollados se han aplicado a problemas de dominios médicos y más recientemente a la bioinformática.

En estas líneas se pretende recoger las publicaciones más relevantes que han surgido como fruto tangible de la investigación llevada a cabo. Dichas publicaciones se agrupan en cinco líneas: redes Bayesianas, clustering, algoritmos genéticos, algoritmos de estimación de distribuciones y aplicaciones.

Por lo que respecta a *redes Bayesianas* los trabajos se han centrado en el aprendizaje estructural de las mismas a partir de datos, tanto en el espacio de los grafos acíclicos dirigidos [2, 4, 25] como en el espacio de órdenes [1]. Se ha investigado asimismo en la búsqueda de la triangulación óptima del grafo moral asociado a una red Bayesiana [3], al igual que en la utilización de las redes Bayesianas como paradigma para representar el comportamiento conjunto de diversos inductores [10].

El problema del *clustering* se ha tratado desde una triple perspectiva. Por una parte en el denominado clustering particional se ha estudiado la influencia de la inicialización en el algoritmo K-means [9]. En relación con el clustering ascendente jerárquico se ha estudiado una aproximación que basándose en los algoritmos genéticos trata el problema como si se tratase de un problema de optimización [6]. Finalmente se ha trabajado en la aproximación probabilista al clustering, efectuándose la misma a partir de redes Bayesianas y redes Gaussianas de distinta complejidad [11, 12, 17, 21], estudiándose asimismo el problema de la reducción de la dimensionalidad [19].

Las publicaciones mas relevantes en el ámbito de los *algoritmos genéticos* se han centrado en un trabajo de recopilación de operadores relacionados con el problema del agente viajero [7], y en el estudio de las condiciones de convergencia de un tipo de algoritmos genéticos elitistas [8].

Los *algoritmos de estimaciones de distribuciones* han sido motivo de investigación durante los últimos años. Fruto del mismo ha sido el libro publicado por Kluwer Academic Publishers [20] de recopilación de las técnicas y métodos que se engloban bajo la denominación de EDA, así como el número especial [26] que en breve verá la luz en la revista International Journal of Approximate Reasoning. Otros trabajos relevantes en esta línea han sido los relacionados con el

estudio de la convergencia de dos tipos de EDAs [14, 24], la aplicación de dichos algoritmos al problema de la selección de variables [13, 16], así como un trabajo de aplicación al problema de macheo inexacto de grafos [22].

Las *aplicaciones* más relevantes las agrupamos en dos áreas: medicina y bioinformática. En cuanto a la primera se han tratado problemas de diagnóstico relacionados con la predicción de la supervivencia en melanoma [5], supervivencia en una unidad de cuidados intensivos [15] y supervivencia en pacientes cirróticos [18]. En el dominio de la bioinformática se ha estudiado el problema de la selección de genes relevantes a partir de información contenida en microarrays de ADN [23].

References

1. P. Larrañaga, C. Kuijpers, R. Murga, Y. Yurramendi (1996). Learning Bayesian network structures by searching for the best ordering with genetic algorithms. *IEEE Transactions on System, Man and Cybernetics*. Vol 26. No. 4, 487-493.
2. P. Larrañaga, M. Poza, Y. Yurramendi, R. Murga, C. Kuijpers (1996). Structure learning of Bayesian networks by genetic algorithms: A performance analysis of control parameters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol. 18. No. 9, 912-926.
3. P. Larrañaga, C. Kuijpers, M. Poza, R. Murga (1997). Decomposing Bayesian networks by genetic algorithms. *Statistics and Computing*. Vol. 7. No. 1, 19-34.
4. R. Etxeberria, P. Larrañaga, J.M. Pikaza (1997). Analysis of the behaviour of the genetic algorithms when searching Bayesian networks from data. *Pattern Recognition Letters*. Vol. 18, No. 11-13, 1269-1273.
5. B. Sierra, P. Larrañaga (1998). Predicting the survival in malignant skin melanoma using Bayesian networks automatically induced by genetic algorithms. An empirical comparison between different approaches. *Artificial Intelligence in Medicine*. Vo. 14. No. 1-2, 215-230.
6. J. A. Lozano, P. Larrañaga (1998). Applying genetic algorithms to search for the best hierarchical clustering of a dataset. *Pattern Recognition Letters* **20**, 911-918.
7. P. Larrañaga, C. Kuijpers, R. Murga, I. Inza, S. Dizdarevich (1999). Evolutionary algorithms for the travelling salesman problem: A review of representations and operators. *Artificial Intelligence Review*, **13**, 129-170.
8. J. A. Lozano, P. Larrañaga, M. Graña, F. X. Albizuri (1999). Genetic algorithms: bridging the convergence gap. *Theoretical Computer Science*. **229**, 11-22.
9. J. M. Peña, J. A. Lozano, P. Larrañaga (1999). An empirical comparison of four initialization methods for the k -means algorithm. *Pattern Recognition Letters*. **20**, 1027-1040.
10. I. Inza, P. Larrañaga, B. Sierra, R. Etxeberria, J. A. Lozano, J. M. Peña (1999). Representing the joint behaviour of machine learning inducers by Bayesian networks. *Pattern Recognition Letters*. **20**, (11-13), 1201-1209.
11. J. M. Peña, J. A. Lozano, P. Larrañaga (1999). Constructive induction for clustering: naive-Bayes models to mixtures of Bayesian networks. *Pattern Recognition Letters*. **20**, (11-13), 1219-1230.
12. J.M. Peña, J.A. Lozano, P. Larrañaga (1999). An improved Bayesian Structural EM algorithm for learning Bayesian networks for clustering. *Pattern Recognition Letters* **21**, 8, 779-786.

13. I. Inza, P. Larrañaga, R. Etxeberria, B. Sierra (2000). Feature Subset Selection by Bayesian networks based optimization. *Artificial Intelligence* **123**, 157-184
14. C. González, J. A. Lozano, P. Larrañaga (2000). Analyzing the PBIL Algorithm by Means of Discrete Dynamical Systems. *Complex Systems*, Vol. 12, No. 4, 465-479.
15. B. Sierra, N. Serrano, P. Larrañaga, E. J. Plasencia, I. Inza, J. J. Jimenez, J. M. De la Rosa, M. J. Mora (2001). Using Bayesian networks in the construction of a bi-level multiclassifier. A case study using intensive care unit patients data. *Artificial Intelligence in Medicine*, **22**, 233-248.
16. I. Inza, P. Larrañaga, B. Sierra (2001). Feature subset selection by Bayesian networks: a comparison with genetic and sequential algorithms. *International Journal of Approximate Reasoning*, **27**, 143-164.
17. J. M. Peña, J. A. Lozano, P. Larrañaga (2001). Performance evaluation of compromise conditional Gaussian networks for data clustering. *International Journal of Approximate Reasoning*, **28**, 23-50.
18. I. Inza, M. Merino, P. Larrañaga, J. Quiroga, B. Sierra, M. Giralá (2001). Feature subset selection by genetic algorithms and estimation of distribution algorithms. A case study in the survival of cirrhotic patients treated with TIPS. *Artificial Intelligence in Medicine*, **23-2**, 187-205.
19. J. M. Peña, J. A. Lozano, P. Larrañaga, I. Inza (2001). Dimensionality reduction in unsupervised learning of conditional Gaussian networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 6, 590-603.
20. P. Larrañaga, J. A. Lozano (2001) (editores). *Estimation of Distribution Algorithms. A New Tool for Evolutionary Computation*. Kluwer Academic Publishers.
21. J. M. Peña, J. A. Lozano, P. Larrañaga (2002). Learning recursive Bayesian multi-nets for clustering by means of constructive induction. *Machine Learning*, **47**, 63-89.
22. E. Bengoetxea, P. Larrañaga, I. Bloch, A. Perchant, C. Boeres (2002). Learning and simulation of Bayesian networks applied to inexact graph matching. *Pattern Recognition*. Aceptado.
23. I. Inza, B. Sierra, R. Blanco, P. Larrañaga (2002). Gene selection by sequential search wrapper approaches in microarray cancer class prediction. *Journal of Intelligent and Fuzzy Systems*. Aceptado.
24. C. González, J.A. Lozano, P. Larrañaga (2002). Mathematical modelling of UMDAc algorithm with tournament selection. Behaviour on linear and quadratic functions. *International Journal of Approximate Reasoning*. Aceptado.
25. R. Blanco, I. Inza, P. Larrañaga (2002). Learning Bayesian networks in the space of structures by estimation of distribution algorithms. *International Journal of Intelligent Systems*. Aceptado.
26. P. Larrañaga, J.A. Lozano (2003). editores invitados del número especial *Synergies Between Probabilistic Graphical Models and Evolutionary Computation* de la revista *International Journal of Approximate Reasoning*.