

1.5. Redes bayesianas usadas como clasificadores.

Dentro del aprendizaje automático se, conoce como *clasificación* a la tarea de asociar los valores que toma un conjunto de variables (también denominadas *características*) con un conjunto discreto de valores, denominados *clases*. La idea es construir de forma automática un modelo, denominado *clasificador*, a partir de un conjunto de datos, y luego, utilizarlo para asignar clases a nuevos datos que no han sido usados en la construcción del modelo.

El funcionamiento de un clasificador depende en gran medida de las propiedades de los datos que queramos clasificar, por lo que no hay un tipo de clasificador que sea mejor que el resto de metodologías para todos los problemas [360].

Dentro del aprendizaje automático para clasificación podemos distinguir entre *clasificación no supervisada* y *clasificación supervisada*. Se diferencian en que el aprendizaje supervisado utiliza un conjunto de muestras previamente clasificadas, es decir, que para cada caso de la muestra conteniendo valores de las variables tiene asociada una clase. En clasificación no supervisada, por el contrario, se tiene el conjunto de datos sin ningún conocimiento *a priori* de su clase: no se conoce la clase de cada caso o incluso, en la mayoría de los casos, se desconoce cuantas clases hay.

1.5.1. Clasificación supervisada y no supervisada.

En *clasificación supervisada* el conjunto de variables a considerar será $\mathbf{V} = \mathbf{X} \cup \{C\}$. Donde C es la variable a clasificar (también denominada variable distinguida o, simplemente, clase) y las variables en \mathbf{X} serán usadas para predecir los valores de C y se les denomina variables predictoras, variables inductoras, características o atributos. El objetivo es describir la variable clase en función de los atributos y poder calcular la probabilidad condicionada de la clase dada cualquier configuración de los atributos, $p(C|\mathbf{X})$. Las redes bayesianas usadas en tareas de clasificación supervisada se denominan genéricamente, *clasificadores bayesianos*.

En *clasificación no supervisada* (o agrupamiento -*clustering*-), partiendo de un conjunto de variables $\mathbf{V} = \mathbf{X}$ se trata de asignar cada caso a un grupo (*clúster*) C , es decir, el objetivo es descubrir una estructura de clases en los

datos, de tal manera que los casos que pertenezcan a una misma clase o grupo presenten una gran homogeneidad, mientras que los casos que pertenezcan a distintos agrupamientos o clasificaciones deben ser muy heterogéneos entre sí. Obsérvese, que se puede ver la clasificación no supervisada, como una clasificación supervisada donde todos los valores de la variable clase son desconocidos, es decir, como aprendizaje de datos incompletos. Dentro de la clasificación no supervisada hay distintos métodos de aprendizaje, usando redes bayesianas, a partir de datos incompletos [67, 115, 335, 231].

1.5.2. Evaluación de clasificadores.

Un detalle importante cuando construimos clasificadores es cuantificar de alguna manera lo buenos o malos que son [350]. Por ejemplo, si usamos un clasificador para detección de cáncer no nos podemos permitir el lujo de que falle tres de cada cuatro pacientes. A la hora de evaluar un clasificador se hará teniendo en cuenta distintos criterios, como puede ser el tiempo que tardamos en construirlo, la interpretabilidad del modelo obtenido, la sencillez del modelo (cuanto más sencillo, mayor capacidad de abstracción) o diferencias respecto al modelo original; pero al que más atención se le presta es a la *precisión* del clasificador (o, a la inversa, la tasa de error) que posee.

La precisión de un clasificador es la probabilidad con la que clasifica correctamente un caso seleccionado al azar [196], o también, lo podemos ver como el número de casos clasificados correctamente entre el número total de elementos.

$$\text{precisión} = \frac{\text{número de aciertos}}{\text{número de casos}}$$

Además de ser la medida más aceptada para la evaluación de un clasificador, la precisión es utilizada en algunos procedimientos para guiar la construcción del clasificador, por ello vamos a exponer distintas formas de obtener su valor:

- *Estimación por resustitución (resubstitution estimate) o error aparente:* este modo, el más simple, consiste en utilizar los mismos datos que se han utilizado para construir el clasificador para ver cuántos predice correctamente.

- *Holdout*: se basa en partir el conjunto de datos aleatoriamente en dos grupos. El denominado conjunto de entrenamiento (normalmente $2/3$ del número total de casos) con el cual se construye el clasificador y el conjunto de testeo o validación (construido con el $1/3$ restante) usado para estimar la precisión del clasificador.
- *Remuestreo (random subsampling)*: es una variación del sistema anterior donde se realizan diferentes particiones de los conjuntos de entrenamiento y test. Obteniéndose la precisión del clasificador a partir de la media obtenida en los distintos conjuntos de test.
- *Validación cruzada de k-hojas (k-fold cross-validation)* [326]: se puede ver como una generalización del criterio de remuestreo. Hacemos k particiones del conjunto de datos mutuamente excluyentes y de igual tamaño. $k - 1$ conjuntos se utilizan para construir el clasificador y se valida con el conjunto restante. Este paso se efectúa k veces y la estimación de la precisión del clasificador se obtiene como la media de las k mediciones realizadas.
- *Dejar-uno-fuera (leave-one-out)* [203]: es un caso particular de la validación cruzada, donde se hacen tantas particiones como casos tenga el conjunto de datos. De esta forma los conjuntos de validación tienen un sólo caso y los de entrenamiento todos los casos menos ese en particular. Al tener que construir el clasificador tantas veces como casos tenga el conjunto de datos se hace bastante costoso en tiempo, a no ser que se pueda obtener una fórmula cerrada para el error.
- *Bootstrapping* [102]: para evaluar un clasificador de forma efectiva cuando se tienen conjuntos de datos con pocas muestras se suele utilizar el método de dejar-uno-fuera. No obstante, con conjuntos pequeños de datos, suele mostrar una varianza alta. El bootstrapping (o bootstrap) es un método de remuestreo propuesto por Bradley Efron en 1979 [101], donde para una muestra de tamaño n se genera el conjunto de entrenamiento con n casos mediante muestreo con reemplazamiento, es decir, cogemos un caso de forma aleatoria del conjunto de datos para el conjunto de entrenamiento y lo volvemos a dejar en el conjunto de datos, de esta forma puede haber casos repetidos en el conjunto de entrenamiento. El conjunto de test se genera cogiendo aquellos casos que no estén en el conjunto de entrenamiento. Este proceso (crear el conjunto de

entrenamiento mediante muestreo con reemplazo) se realiza un número elevado de veces. La media de las distintas precisiones calculadas sirve como estimación de la precisión verdadera.

Algunas veces es interesante no sólo conocer la precisión del clasificador o la tasa de error, sino que es importante el sentido en el que se equivoca. Pongamos un ejemplo extremo: estamos diagnosticando mediante un clasificador si un paciente tiene o no cáncer; si el clasificador se equivoca y toma a una persona sana como enferma (ésto se conoce como *falso positivo*), es un error que probablemente se descubrirá más tarde en posteriores análisis (en el peor de los casos se habrá generado una serie de pruebas y preocupaciones innecesarias). Por el contrario, si el clasificador toma a una persona como sana cuando en realidad está enfermo (lo que se conoce como un *falso negativo*), quizás a la hora de solucionar el error ya sea demasiado tarde (perdemos la posibilidad de tratar al paciente a tiempo).

Cuando distinguir entre los distintos tipos de errores es importante, entonces se puede usar una *matriz de confusión* (también llamada *tabla de contingencia*) para mostrar los diferentes tipos de error. Si tenemos un problema con dos clases (por ejemplo, sano y enfermo), como se puede ver en la tabla 1.1, un clasificador puede dar la siguiente salida para un nuevo caso: verdadero positivo, predice que el paciente tiene la enfermedad y es verdad, verdadero negativo si acierta que el paciente no está enfermo, falso positivo pronostica enfermo estando el paciente sano y, finalmente, falso negativo, predice sano pero el paciente está enfermo.

	Enfermo	Sano
Enfermo	Verdadero positivo	Falso positivo
Sano	Falso negativo	Verdadero negativo

Tabla 1.1: Matriz de confusión para un problema con dos clases.

A partir de la matriz de confusión podemos construir algunas medidas que nos serán de utilidad. La *sensibilidad* es la probabilidad de clasificar correctamente a un individuo enfermo, por tanto, es la capacidad del clasificador para detectar la clase positiva (la que más interesa, en el ejemplo, que

está enfermo), se define como:

$$\text{sensibilidad} = \frac{\text{verdaderos_positivos}}{\text{verdaderos_positivos} + \text{falsos_negativos}}$$

por otro lado, definimos la *especificidad* como la probabilidad de clasificar correctamente a un individuo sano, en otras palabras, se puede definir la especificidad como la capacidad de clasificar correctamente a un individuo cuyo estado real sea negativo para la prueba que se hace (en nuestro ejemplo, que esté sano). Se puede calcular a partir de la matriz de confusión como:

$$\text{especificidad} = \frac{\text{verdaderos_negativos}}{\text{verdadero_negativos} + \text{falsos_positivos}}$$

En problemas donde distinguir el tipo de error es importante, se pueden utilizar las *curvas ROC* (del inglés, *Receiver Operating Characteristics*) [110]. Las curvas ROC aunque tienen su origen en la detección de señales de radar, se usan habitualmente en la toma de decisiones médicas [330] y además nos van a permitir evaluar de forma gráfica el funcionamiento de un clasificador. Son curvas en las que se presenta la sensibilidad en función de los falsos positivos (complementario de la especificidad) para distintos resultados de un clasificador.

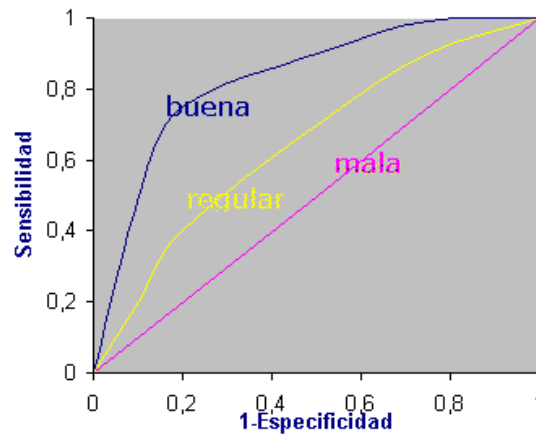


Figura 1.5: Tipos de curvas ROC.

Cuanto mayor sea el área bajo la curva ROC, mejor será el clasificador. La mejor predicción posible sería un método que pasara por la esquina superior izquierda, ya que representaría un 100 % de sensibilidad (no hay falsos negativos) y un 100 % de especificidad (no hay falsos positivos). Por tanto, cuanto más distante está la curva de la diagonal mejor será el clasificador, como se puede ver en la figura 1.5.

1.5.3. Enfoques de filtrado y de envoltura en clasificadores.

La construcción de un clasificador, en algunos casos, conlleva un proceso de búsqueda en la estructura del mismo, con el objetivo de maximizar la precisión del clasificador final y así, obtener un mejor modelo. Esa búsqueda de la estructura de los clasificadores pueden hacerse en base a dos enfoques:

- Filtrado (*filter*): se utiliza una medida (por ejemplo, la verosimilitud, la entropía o la información mutua) para medir la bondad de las distintas estructuras candidatas en el proceso de búsqueda. Esta medida es independiente del clasificador utilizado.
- Envoltura (*wrapper*): se realiza una búsqueda guiada por la precisión del clasificador con el que se está trabajando, como si fuera una caja negra que evalúa cada una de las estructuras candidatas del espacio de búsqueda. Normalmente se utiliza una validación cruzada sobre cada una de las estructuras candidatas, por lo que es computacionalmente un método bastante costoso.

Los métodos de envoltura presentan un mejor comportamiento en la precisión de los clasificadores que los métodos de filtrado, algo lógico pues optimizan la precisión del clasificador en su proceso de búsqueda. No obstante los clasificadores de filtrado suelen presentar un comportamiento más robusto frente a un sobreajuste a los datos y son más eficientes.

1.5.4. Clasificadores bayesianos.

En este trabajo nos vamos a centrar en el uso de clasificadores bayesianos, es decir, redes bayesianas utilizadas para tareas de clasificación supervisada

- [93] A. P. DEMPSTER, N. M. LAIRD Y D. B. RUBIN. Maximum-likelihood from incomplete data via the EM algorithm. *Journal of Royal Statistical Society B* **39**, 1–38 (1977).
- [94] J. L. DERISI, V. R. IYER Y P. O. BROWN. Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* **278**(5338), 680–686 (1997).
- [95] Q. DIAO, W. HU, H. ZHONG, J. LI, F. XUE, T. WANG Y Y. ZHAN. Disease gene explorer: Display disease gene dependency by combining Bayesian networks with clustering. En “3rd International IEEE Computer Society Computational Systems Bioinformatics Conference”, págs. 574–575. IEEE Computer Society (2004).
- [96] F. J. DÍEZ. Aplicaciones de los modelos gráficos probabilísticos en medicina. En J. A. GÁMEZ Y J. M. PUERTA, editores, “Sistemas Expertos Probabilísticos”, Ciencia y Técnica num. 20, págs. 239–263. Universidad de Castilla-La Mancha, 1 ed. (1998).
- [97] C. DING Y H. PENG. Minimum redundancy feature selection from microarray gene expression data. *Journal of Bioinformatics and Computational Biology* **3**(2), 185–205 (2005).
- [98] A. DJEBBARI Y J. QUACKENBUSH. Seeded Bayesian networks: constructing genetic networks from microarray data. *BMC Systems Biology* **2**, 57 (2008).
- [99] J. DOUGHERTY, R. KOHAVI Y M. SAHAMI. Supervised and unsupervised discretization of continuous features. En “International Conference on Machine Learning”, págs. 194–202 (1995).
- [100] R. DUDA Y P. HART. “Pattern classification and scene analysis”. John Wiley and Sons, New York (1973).
- [101] B. EFRON. Bootstrap methods: Another look at the jackknife. *The Annals of Statistics* **7**, 1–26 (1979).
- [102] B. EFRON Y R. TIBSHIRANI. “An Introduction to the bootstrap”. Chapman and Hall (1993).

- [103] M. B. EISEN, P. T. SPELLMAN, P. O. BROWN Y D. BOTSTEIN. Cluster analysis and display of genome-wide expression patterns. *Proceedings of the National Academy of Sciences of the United States of America (PNAS)* **95**(25), 14863–8 (1998).
- [104] G. ELIDAN Y N. FRIEDMAN. Learning hidden variable networks: The information bottleneck approach. *Journal of Machine Learning Research* **6**, 81–127 (2005).
- [105] ELVIRA CONSORTIUM. Elvira: An environment for probabilistic graphical models. En J. GÁMEZ Y A. SALMERÓN, editores, “Procs. of the 1st European Workshop on Probabilistics Graphical Models (PGM 2002)”, págs. 222–230 (2002).
- [106] S. B. ENGLISH, S. SHIH, M. F. RAMONI, L. E. SMITH Y A. J. BUTTE. Use of Bayesian networks to probabilistically model and improve the likelihood of validation of microarray findings by RT-PCR. *Journal of Biomedical Informatics* **42**(2), 287–295 (2009).
- [107] F. ESPOSITO, D. MALERBA Y G. SEMERARO. A comparative analysis of methods for pruning decision trees. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **19**(5), 476–491 (1997).
- [108] J. A. FALCONER, B. J. NAUGHTON, D. D. DUNLOP, E. J. ROTH Y D. STRASSER. Predicting stroke inpatient rehabilitation outcome using a classification tree approach. *Archives of Physical Medicine and Rehabilitation* **75**(6), 619 (1994).
- [109] L. FAN, K.-L. POHA Y P. ZHOUB. A sequential feature extraction approach for naïve bayes classification of microarray data. *Expert Systems with Applications* **36**(6), 9919–9923 (2009).
- [110] T. FAWCETT. ROC graphs: Notes and practical considerations for researchers. Informe técnico HPL-2003-4, HP Labs, Palo Alto, USA (2004).
- [111] U. FAYYAD Y K. IRANI. Multi-valued interval discretization of continuous-valued attributes for classification learning. En “Proceeding of the 13th International joint Conference on Artificial Intelligence”, págs. 1022–1027. Morgan Kaufmann (1993).

- [112] G. F.COOPER. The computational complexity of probabilistic inference using Bayesian belief networks. *Artificial Intelligence* **42**, 393–405 (1990).
- [113] P. E. FILE, P. I. DUGARD Y A. S. HOUSTON. Evaluation of the use of induction in the development of a medical expert system. *Computers and Biomedical Research* **27**(5), 383–395 (1994).
- [114] E. FIX Y J. J. L. HODGES. Discriminatory analysis: nonparametric discrimination: consistency properties. Informe técnico Project 21-49-004, Report Number 4, USAF School of Aviation Medicine (1951).
- [115] N. FRIEDMAN. The Bayesian structural EM algorithm. En G. F. COOPER Y S. MORAL, editores, “Proceedings of the 14th Annual Conference on Uncertainty in Artificial Intelligence (UAI-98)”, págs. 129–138. Morgan Kaufmann Publishers (1998).
- [116] N. FRIEDMAN. Inferring cellular networks using probabilistic graphical models. *Science* **303**(5659), 799–805 (2004).
- [117] N. FRIEDMAN, I. NACHMAN Y D. PEÉR. Learning Bayesian network structure from massive datasets: The ”Sparse Candidate” algorithm. En “UAI ’99: Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence”, págs. 206–215 (1999).
- [118] N. FRIEDMAN, D. GEIGER Y M. GOLDSZMIDT. Bayesian networks classifiers. *Machine Learning* **29**, 131–163 (1997).
- [119] N. FRIEDMAN Y M. GOLDSZMIDT. Learning Bayesian networks with local structure. En E. HORVITZ Y F. V. JENSEN, editores, “Proceedings of the 12th Annual Conference on Uncertainty in Artificial Intelligence (UAI-96)”, págs. 252–262. Morgan Kaufmann Publishers (1996).
- [120] N. FRIEDMAN, M. GOLDSZMIDT Y A. J. WYNER. Data analysis with Bayesian networks: A bootstrap approach. En “Proceedings of the Fifteenth Annual Conference on Uncertainty in Artificial Intelligence (UAI-99)”, págs. 206–215. Morgan Kaufmann (1999).
- [121] N. FRIEDMAN Y D. KOLLER. Being Bayesian about network structure. En “UAI ’00: Proceedings of the 16th Conference on Uncertainty in

- Artificial Intelligence”, págs. 201–210. Morgan Kaufmann Publishers Inc. (2000).
- [122] N. FRIEDMAN, M. LINIAL, I. NACHMAN Y D. PE’ER. Using Bayesian networks to analyze expression data. *Journal of Computational Biology* **7**(3-4), 601–620 (2000).
 - [123] G. GAMBERONI, E. LAMMA, F. RIGUZZI, S. STORARI Y S. VOLINIA. Bayesian networks learning for gene expression datasets. En A. F. FAMILI, J. KOK, J. M. PEÑA, A. SIEBES Y A. J. FEELDERS, editores, “Advances in Intelligent Data Analysis VI, 6th International Symposium on Intelligent Data Analysis, IDA 2005, Proceedings”, vol. 3646 de “Lecture Notes in Computer Science”, págs. 109–120. Springer (2005).
 - [124] J. A. GÁMEZ. “Inferencia abductiva en redes causales”. Tesis Doctoral, Departamento de Ciencias de la Computación e I.A. Escuela Técnica Superior de Ingeniería Informática. Universidad de Granada (1998).
 - [125] A. P. GASCH, P. T. SPELLMAN, C. M. KAO, O. C. HAREL, M. B. EISEN, G. STORZ, D. BOTSTEIN Y P. O. BROWN. Genomic expression programs in the response of yeast cells to environmental changes. *Molecular Biology of the Cell* **11**(12), 4241–4257 (2000).
 - [126] D. GEIGER Y D. HECKERMAN. Learning Gaussian networks. En R. L. DE MÁNTARAS Y D. POOLE, editores, “UAI ’94: Proceedings of the Tenth Annual Conference on Uncertainty in Artificial Intelligence”, págs. 235–243. Morgan Kaufmann Publishers (1994).
 - [127] D. GEIGER Y D. HECKERMAN. A characterization of the Dirichlet distribution with application to learning Bayesian networks. En P. BESNARD Y S. HANKS, editores, “UAI ’95: Proceedings of the Eleventh Annual Conference on Uncertainty in Artificial Intelligence”, págs. 196–207. Morgan Kaufmann (1995).
 - [128] D. GEIGER Y D. HECKERMAN. Knowledge representation and inference in similarity networks and Bayesian multinets. *Artificial Intelligence* **82**, 45–74 (1996).

- [129] D. GEIGER Y D. HECKERMAN. A characterization of the Dirichlet distribution through global and local parameter independence. *Annals of Statistics* **25**, 1344–1369 (1997).
- [130] D. GEIGER, A. PAZ Y J. PEARL. Learning simple causal structures. *Knowledge acquisition as modeling* **8**, 231–247 (1993).
- [131] O. GEVAERT, F. D. SMET, D. TIMMERMAN, Y. MOREAU Y B. D. MOOR. Predicting the prognosis of breast cancer by integrating clinical and microarray data with Bayesian networks. *Bioinformatics* **22**(14), e184–e190 (2006).
- [132] O. GEVAERT, S. VAN VOOREN Y B. DE MOOR. Integration of microarray and textual data improves the prognosis prediction of breast, lung and ovarian cancer patients. *Pacific Symposium on Biocomputing* págs. 279–290 (2008).
- [133] Z. GHAHRAMANI. Learning dynamic Bayesian networks. En “Adaptive Processing of Sequences and Data Structures, International Summer School on Neural Networks”, vol. 1387 de “Lecture Notes in Computer Science”, págs. 168–197. Springer-Verlag (1998).
- [134] D. K. GIFFORD. Blazing pathways through genetic mountains. *Science* **293**, 2049–2051 (2001).
- [135] S. B. GILLISPIE Y C. LEMIEUX. Enumerating Markov equivalence classes of acyclic digraph models. En “UAI ’01: Proceedings of the 17th Conference in Uncertainty in Artificial Intelligence”, págs. 171–177. Morgan Kaufmann Publishers Inc. (2001).
- [136] T. R. GOLUB, D. K. SLONIM, P. TAMAYO, C. HUARD, M. GAASENBEEK, J. P. MESIROV, H. COLLIER, M. L. LOH, J. R. DOWNING, M. A. CALIGIURI, C. D. BLOOMFIELD Y E. S. LANDER. Molecular classification of cancer: class discovery and class prediction by gene expression monitoring. *Science* **286**, 531–537 (1999).
- [137] I. J. GOOD. “The estimation of probabilities”. The MIT Press (1965).
- [138] T. GRAEPEL, M. BURGER Y K. OBERMAYER. Self-organizing maps: Generalizations and new optimization techniques. *Neurocomputing* **20**, 173–190 (1998).

- [139] A. J. F. GRIFFITHS, J. H. MILLER, D. T. SUZUKI, R. C. LEWONTIN Y W. M. GELBART. “Genética”. Interamericana, España, 7 ed. (2002).
- [140] M. GRZEGORCZYK, D. HUSMEIER, K. D. EDWARDS, P. GHAZAL Y A. J. MILLAR. Modelling non-stationary gene regulatory processes with a non-homogeneous Bayesian network and the allocation sampler. *Bioinformatics* **24**(18), 2071–2078 (2008).
- [141] M. A. HALL Y L. A. SMITH. Feature subset selection: a correlation based filter approach. En “International Conference on Neural Information Processing and Intelligent Information Systems”, págs. 855–858. Springer (1997).
- [142] A. J. HARTEMINK, D. K. GIFFORD, T. S. JAAKKOLA Y R. A. YOUNG. Using graphical models and genomic expression data to statistically validate models of genetic regulatory networks. *Pacific Symposium on Biocomputing* págs. 422–33 (2001).
- [143] A. J. HARTEMINK, D. K. GIFFORD, T. S. JAAKKOLA Y R. A. YOUNG. Bayesian methods for elucidating genetic regulatory networks. *IEEE Intelligent Systems* **17**(2), 37–43 (2002).
- [144] A. J. HARTEMINK, D. K. GIFFORD, T. S. JAAKKOLA Y R. A. YOUNG. Combining location and expression data for principled discovery of genetic regulatory network models. *Pacific Symposium on Biocomputing* págs. 437–49 (2002).
- [145] T. J. HASTIE Y R. J. TIBSHIRANI. “Generalized additive models”. London: Chapman & Hall (1990).
- [146] S. HAUTANIEMI, H. EDGREN, P. VESANEN, M. WOLF, A. K. JÄRVINEN, O. YLI-HARJA, J. ASTOLA, O. P. KALLIONIEMI Y O. MONNI. A novel strategy for microarray quality control using Bayesian networks. *Bioinformatics* **19**(16), 2031–2038 (2003).
- [147] D. HECKERMAN. “Probabilistic similarity networks”. ACM Doctoral dissertation award series, MIT Press (1991).

- [148] D. HECKERMAN. A tutorial on learning with Bayesian networks. Informe técnico MSR-TR-95-06, Microsoft Research, Advanced Technology Division (1995).
- [149] D. HECKERMAN. Bayesian networks for knowledge discovery. *Advances in knowledge discovery and data mining* págs. 273–305 (1996).
- [150] D. HECKERMAN, D. GEIGER Y D. M. CHICKERING. Learning Bayesian networks: The combination of knowledge and statistical data. *Machine Learning* **20**(3), 197–243 (1995).
- [151] I. HEDENFALK, Y. C. D. DUGGAN, M. RADMACHER, M. BITTNER, R. SIMON, P. MELTZER, B. GUSTERSON, M. ESTELLER, O. KALLIONIEMI, B. WILFOND, A. BORG Y J. TRENT. Gene-expression profiles in hereditary breast cancer. *The new england journal of medicine* **344**(9), 539–548 (2001).
- [152] P. HELMAN, R. VEROFF, S. R. ATLAS Y C. WILLMAN. A Bayesian network classification methodology for gene expression data. *Journal of Computational Biology* **11**(4), 581 – 615 (2004).
- [153] L. D. HERNÁNDEZ. Algoritmos de propagación i: Métodos exactos. En J. A. GÁMEZ Y J. M. PUERTA, editores, “Sistemas Expertos Probabilísticos”, Ciencia y Técnica num. 20, págs. 41–64. Universidad de Castilla-La Mancha, 1 ed. (1998).
- [154] E. HERSKOVITS Y G. F. COOPER. Kutató: An entropy-driven system for the construction of probabilistic expert systems from databases. En “Proceedings of the Sixth Conference on Uncertainty in Artificial Intelligence”, págs. 54–62 (1990).
- [155] A. HOPKINS, J. JOHNSON, S. LAHART, D. McLAUGHLIN, C. W. WARNER, M. Q. WRIGHT Y J. D. MATON. “Cells, building blocks of life”. Pearson Prentice Hall (1997).
- [156] K. HORNIK, C. BUCHTA Y A. ZEILEIS. Open-source machine learning: R meets Weka. *Computational Statistics* **24**(2), 225–232 (2009).
- [157] Y. HUANG, J. WANG, J. ZHANG, M. SANCHEZ Y Y. WANG. Bayesian inference of genetic regulatory networks from time series microarray data using dynamic Bayesian networks. *Journal of Multimedia* **2**(3), 46–56 (2007).

- [158] Z. HUANG, J. LI, H. SU, G. S. WATTS Y H. CHEN. Large-scale regulatory network analysis from microarray data: Modified Bayesian network learning and association rule mining. *Decision Support Systems* **43**(4), 1207–1225 (2007).
- [159] T. R. HUGHES, M. J. MARTON, A. R. JONES, C. J. ROBERTS, R. STOUGHTON, C. D. ARMOUR, H. A. BENNETT, E. COFFEY, H. DAI, Y. D. HE, M. J. KIDD, A. M. KING, M. R. MEYER, D. SLADE, P. Y. LUM, S. B. STEPANIANTS, D. D. SHOEMAKER, D. GACHOTTE, K. CHAKRABURTTY, J. SIMON, M. BARD Y S. H. FRIEND. Functional discovery via a compendium of expression profiles. *Cell* **102**(1), 109–126 (2000).
- [160] E. HUN, J. MARIN Y P. STONE. Experiments in induction. *Academic Press* (1966).
- [161] D. HUSMEIER. Reverse engineering of genetic networks with Bayesian networks. *Biochemical Society Transactions* **31**(Pt 6), 1516–8 (2003).
- [162] D. HUSMEIER. Sensitivity and specificity of inferring genetic regulatory interactions from microarray experiments with dynamic Bayesian networks. *Bioinformatics* **19**(17), 2271–82 (2003).
- [163] K. B. HWANG, D. Y. CHO, S. W. PARK, S. D. KIM Y B. T. ZHANG. Applying machine learning techniques to analysis of gene expression data: Cancer diagnosis. En S. M. LIN Y K. F. JOHNSON, editores, “Methods of Microarray Data Analysis (Proceedings of CAMDA’00),”, págs. 167–182. Kluwer Academic Publishers (2002).
- [164] K. B. HWANG Y B. T. ZHANG. Bayesian model averaging of Bayesian network classifiers over multiple node-orders: application to sparse datasets. *IEEE Transactions on Systems, Man, and Cybernetics, Part B* **35**(6), 1302–1310 (2005).
- [165] J. G. IBRAHIM, M. H. CHEN Y R. J. GRAY. Bayesian models for gene expression with DNA microarray data. *Journal of the American Statistical Association* **97**, 88–99 (2002).
- [166] S. IMOTO, T. GOTO Y S. MIYANO. Estimation of genetic networks and functional structures between genes by using Bayesian network and

- nonparametric regression. En “Pacific Symposium on Biocomputing”, vol. 7, págs. 175–186 (2002).
- [167] S. IMOTO, T. HIGUCHI, T. GOTO Y S. MIYANO. Error tolerant model for incorporating biological knowledge with expression data in estimating. *Statistical Methodology* **3**(1), 1–16 (2006).
- [168] S. IMOTO, T. HIGUCHI, T. GOTO, K. TASHIRO, S. KUHARA Y S. MIYANO. Combining microarrays and biological knowledge for estimating gene networks via Bayesian networks. En “2nd IEEE Computer Society Bioinformatics Conference (CSB 2003)”, págs. 104–113. IEEE Computer Society (2003).
- [169] S. IMOTO, T. HIGUCHI, T. GOTO, K. TASHIRO, S. KUHARA Y S. MIYANO. Combining microarrays and biological knowledge for estimating gene networks via Bayesian networks. *Journal of Bioinformatics and Computational Biology* **2**(1), 77–98 (2004).
- [170] S. IMOTO, S. KIM, T. GOTO, S. MIYANO, S. ABURATANI, K. TASHIRO Y S. KUHARA. Bayesian network and nonparametric heteroscedastic regression for nonlinear modeling of genetic network. *Journal of Bioinformatics and Computational Biology* **1**(2), 231–52 (2003).
- [171] S. IMOTO, H. SHIMODAIRA, S. KIM, S. ABURATANI, K. TASHIRO, S. KUHARA Y S. MIYANO. Bootstrap analysis of gene networks based on Bayesian networks and nonparametric regression. En “Proceedings of the International Conference on Genome Informatics”, vol. 13, págs. 369–370 (2002).
- [172] S. IMOTO, K. SUNYONG, T. GOTO, S. ABURATANI, K. TASHIRO, S. KUHARA Y S. MIYANO. Bayesian network and nonparametric heteroscedastic regression for nonlinear modeling of genetic network. En “CSB ’02: Proceedings of the IEEE Computer Society Conference on Bioinformatics”, pág. 219. IEEE Computer Society (2002).
- [173] I. INZA, P. LARRAÑAGA, R. BLANCO Y A. J. CERROLAZA. Filter versus wrapper gene selection approaches in DNA microarray domains. *Artificial Intelligence in Medicine* **31**(2), 91–103 (2004).

- [174] I. INZA, P. LARRAÑAGA, R. ETXEBERRIA Y B. SIERRA. Feature subset selection by Bayesian network-based optimization. *Artificial Intelligence* **123**(1-2), 157–184 (2000).
- [175] I. INZA, B. SIERRA, R. BLANCO Y P. LARRAÑAGA. Gene selection by sequential search wrapper approaches in microarray cancer class prediction. *Journal of Intelligent and Fuzzy Systems* **12**(1), 25–34 (2002).
- [176] G. H. JOHN, R. KOHAVI Y K. PFLEGER. Irrelevant features and the subset selection problem. En “International Conference on Machine Learning”, págs. 121–129 (1994).
- [177] G. JUDMAIER, P. MEYERSBACH, G. WEISS, H. WACHTER Y G. REIBNEGGER. The role of neopterin in assessing disease activity in Crohn’s disease: Classification and regression trees. *The American Journal of Gastroenterology* **88**, 706–711 (1993).
- [178] T. KAMIMURA, H. SHIMODAIRA, S. IMOTO, S. KIM, K. TASHIRO, S. KUHARA Y S. MIYANO. Multiscale bootstrap analysis of gene networks based on Bayesian networks and nonparametric regression. En “Proceedings of the International Conference on Genome Informatics”, vol. 14, págs. 350–351 (2002).
- [179] M. KANEHISA, M. ARAKI, S. GOTO, M. HATTORI, M. HIRAKAWA, M. ITOH, T. KATAYAMA, S. KAWASHIMA, S. OKUDA, T. TOKIMATSU Y Y. YAMANISHI. KEGG for linking genomes to life and the environment. *Nucleic Acids Research* **36**, 480–484 (2008).
- [180] S. A. KAUFFMAN. Metabolic stability and epigenesis in randomly constructed genetic nets. *Journal of Theoretical Biology* **22**(3), 437–467 (1969).
- [181] E. J. KEOGH Y M. J. PAZZANI. Learning augmented Bayesian classifiers: A comparison of distribution-based and classification-based approaches. En “Proceedings of 7th International Workshop on Artificial Intelligence and Statistics”, págs. 225–230 (1999).
- [182] E. J. KEOGH Y M. J. PAZZANI. Learning the structure of augmented Bayesian classifiers. *International Journal on Artificial Intelligence Tools* **11**(4), 587–601 (2002).

- [183] J. KHAN, J. S. WEI, M. RINGNÉR, L. H. SAAL, M. LADANYI, F. WESTERMANN, F. BERTHOLD, M. SCHWAB, C. R. ANTONESCU, C. PETERSON Y P. S. MELTZER. Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. *Nature Medicine* **7**, 673–679 (2001).
- [184] A. B. KHODURSKY, B. J. PETER, N. R. COZZARELLI, D. BOTSSTEIN, P. O. BROWN Y C. YANOFSKY. DNA microarray analysis of gene expression in response to physiological and genetic changes that affect tryptophan metabolism in *Escherichia coli*. *Proceeding National Academy of Sciences of USA* **97**(22), 12170–5 (2000).
- [185] H. KIM, G. H. GOLUB Y H. PARK. Missing value estimation for DNA microarray gene expression data: local least squares imputation. *Bioinformatics* **21**(2), 187–198 (2005).
- [186] S. KIM, S. IMOTO Y S. MIYANO. Dynamic Bayesian network and non-parametric regression model for inferring gene networks. En “Proceedings of the International Conference on Genome Informatics”, vol. 13, págs. 371–372 (2002).
- [187] S. KIM, S. IMOTO Y S. MIYANO. Dynamic Bayesian network and nonparametric regression for nonlinear modeling of gene networks from time series gene expression data. En C. PRIAMI, editor, “Computational Methods in Systems Biology, First International Workshop, CMSB 2003”, vol. 2602 de “Lecture Notes in Computer Science”, págs. 104–113. Springer (2003).
- [188] S. KIM, S. IMOTO Y S. MIYANO. Inferring gene networks from time series microarray data using dynamic Bayesian networks. *Briefings in Bioinformatics* **4**(3), 228–35 (2003).
- [189] S. KIM, S. IMOTO Y S. MIYANO. Dynamic Bayesian network and nonparametric regression for nonlinear modeling of gene networks from time series gene expression data. *Biosystems* **75**(1-3), 57–65 (2004).
- [190] L. KLEBANOV Y A. YAKOVLEV. How high is the level of technical noise in microarray data? *Biology Direct* **2**, 2–9 (2007).
- [191] S. KNUDSEN. “Cancer diagnostics with DNA microarrays”. John Wiley & Sons, Inc. (2006).

- [192] Y. KO, C. ZHAI Y S. L. RODRIGUEZ-ZAS. Inference of gene pathways using Gaussian mixture models. En “BIBM ’07: Proceedings of the 2007 IEEE International Conference on Bioinformatics and Biomedicine”, págs. 362–367. IEEE Computer Society (2007).
- [193] T. KOCKA Y R. CASTELO. Improved learning of Bayesian networks. En “Proceedings of the 17th Annual Conference on Uncertainty in Artificial Intelligence (UAI-01)”, págs. 269–276. Morgan Kaufmann Publishers (2001).
- [194] R. KOHAVI. A study of cross-validation and bootstrap for accuracy estimation and model selection. En “Fourteenth International Joint Conference on Artificial Intelligence”, págs. 1137–1145 (1995).
- [195] R. KOHAVI. Scaling up the accuracy o naïve-bayes classifier: A decision tree hybrid. *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*. (1996).
- [196] R. KOHAVI. “Wrappers for performance enhancement and oblivious decision graphs”. Tesis Doctoral, Stanford University, Stanford, CA, USA (1996).
- [197] R. KOHAVI Y G. H. JOHN. Wrappers for feature subset selection. *Artificial Intelligence* **97**(1-2), 273–324 (1997).
- [198] R. KOHAVI, D. SOMMERFIELD Y J. DOUGHERTY. Data mining using MLC++: A machine learning library in C++. *International Journal on Artificial Intelligence Tools* **6**(4), 537–566 (1997).
- [199] P. KOKOL, M. MERNIK, J. ZAVRSNIK Y K. KANCLER. Decision trees based on automatic learning and their use in cardiology. *Journal of Medical Systems* **18**(4), 201 (1994).
- [200] A. N. KOLMOGOROV. “Foundations of the theory of probability”. Chelsea Publishing Company, New York, 2 ed. (1956).
- [201] I. KONONENKO. Semi-naïve Bayesian classifier. En “European Working Session on Learning on Machine Learning”, págs. 206–219 (1991).
- [202] I. KONONENKO. Inductive and Bayesian learning in medical diagnosis. *Applied Artificial Intelligence* **4**(7), 317–337 (1993).

- [203] P. LACHENBRUCH Y A. MICKEY. Estimation of error rates in discriminant analysis. *Technometrics* **10**, 1–11 (1968).
- [204] W. LAM Y F. BACCHUS. Learning Bayesian belief networks: An approach based on the MDL principle. *Computational Intelligence* **10**, 269–294 (1994).
- [205] P. LANGELY Y S. SAGE. Induction of selective Bayesian classifiers. *Proceeding of the Tenth Conference on Uncertainty in Artificial Intelligence* págs. 399–406 (1998).
- [206] P. LANGLEY, W. IBA Y K. THOMPSON. An analysis of Bayesian classifiers. En “National Conference on Artificial Intelligence”, págs. 223–228 (1992).
- [207] P. LARRAÑAGA, C. M. H. KUIJPERS, R. H. MURGA Y Y. YURRAMENDI. Learning Bayesian network structures by searching for the best ordering with genetic algorithms. *IEEE Transactions on Systems, Man and Cybernetics* **26**, 487–493 (1996).
- [208] P. LARRAÑAGA Y J. A. LOZANO. “Estimation of distribution algorithms: A new tool for evolutionary computation”. Genetic Algorithms and Evolutionary Computation. Kluwer Academic Publishers (2001).
- [209] P. LARRAÑAGA, M. POZA, Y. YURRAMENDI, R. H. MURGA Y C. M. H. KUIJPERS. Structure learning of Bayesian networks by genetic algorithms: A performance analysis of control parameters. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **18**(9), 912–926 (1996).
- [210] P. LARSEN, E. ALMASRI, G. CHEN Y Y. DAI. A statistical method to incorporate biological knowledge for generating testable novel gene regulatory interactions from microarray experiments. *BMC Bioinformatics* **8**, 317 (2007).
- [211] S. L. LAURITZEN. Propagation of probabilities, means and variances in mixed graphical association models. *Journal of the American Statistical Association* **87**, 1098–1108 (1992).
- [212] S. L. LAURITZEN Y D. J. SPIEGELHALTER. Local computations with probabilities on graphical structures and their application to expert

- systems (with discussion). *Journal of the Royal Statistical Society, Series B* **50**, 157–224 (1988).
- [213] S. L. LAURITZEN Y N. WERMUTH. Graphical models for associations between variables, some of which are qualitative and some quantitative. *Annals of Statistics* **17**(1), 31–57 (1998).
- [214] P. LE PHILLIP, A. BAHL Y L. H. UNGAR. Using prior knowledge to improve genetic network reconstruction from microarray data. *In Silico Biology* **4**(3), 335–353 (2004).
- [215] R. D. LECLERC. Survival of the sparsest: robust gene networks are parsimonious. *Molecular Systems Biology* **4**(213) (2008).
- [216] T. I. LEE, N. J. RINALDI, F. ROBERT, D. T. ODOM, Z. BAR-JOSEPH, G. K. GERBER, N. M. HANNETT, C. T. HARBISON, C. M. THOMPSON, I. SIMON, J. ZEITLINGER, E. G. JENNINGS, H. L. MURRAY, D. B. GORDON, B. REN, J. J. WYRICK, J.-B. TAGNE, T. L. VOLKERT, E. FRAENKEL, D. K. GIFFORD Y R. A. YOUNG. Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* **298**(5594), 799–804 (2002).
- [217] J. LI, H. LIU Y L. WONG. Mean-entropy discretized features are effective for classifying high-dimensional biomedical data. En M. J. ZAKI, J. T.-L. WANG Y H. TOIVONEN, editores, “Proceedings of the 3rd ACM SIGKDD Workshop on Data Mining in Bioinformatics (BIOKDD 2003)”, págs. 17–24 (2003).
- [218] D. V. LINDLEY. “Introduction to probability and statistics from a Bayesian viewpoint”. Cambridge University Press (1965).
- [219] H. LIU Y M. HIROSHI. “Feature selection for knowledge discovery and data mining”. Kluwer Academic Publishers (1998).
- [220] A. D. LONG, H. J. MANGALAM, B. Y. P. CHAN, L. TOLLERI, G. W. HATFIELD Y P. BALDI. Improved statistical inference from DNA microarray data using analysis of variance and a Bayesian statistical framework. Analysis of global gene expression in *Escherichia coli* K12. *Journal of Biological Chemistry* **276**(23), 19937–19944 (2001).

- [221] M. LÓPEZ, P. MALLORQUÍN Y M. VEGA. “Microarrays y biochips de ADN. Informe de Vigilancia Tecnológica.” Genoma España / CIBT-FGUAM (2002).
- [222] P. J. F. LUCAS. Restricted Bayesian network structure learning. En J. GÁMEZ Y A. SALMERÓN, editores, “Procs. of the 1st European Workshop on Probabilistic Graphical Models (PGM 2002)”, págs. 117–126 (2002).
- [223] P. J. F. LUCAS, L. C. VAN DER GAAG Y A. ABU-HANNA. Bayesian networks in biomedicine and health-care. *Artificial Intelligence in Medicine* **30**(3), 201–214 (2004).
- [224] D. MADIGAN, S. ANDERSSON, M. PERLMAN Y C. VOLINSKY. Bayesian model averaging and model selection for Markov equivalence classes of acyclic digraphs. En “Communications in Statistics: Theory and Methods”, págs. 2493–2519 (1996).
- [225] D. MADIGAN Y A. E. RAFTERY. Model selection and accounting for model uncertainty in graphical models using Occam’s window. *Journal of the American Statistical Association* **89**(428), 1535–1546 (1994).
- [226] F. MARKOWETZ Y R. SPANG. Inferring cellular networks - a review. *BMC Bioinformatics* **8**(6) (2007).
- [227] A. R. MASEGOSA. “Model of Supervised Classification: Applications to Genomics and Information Retrieval.” Tesis Doctoral, Granada University (2009).
- [228] H. H. MCADAMS Y A. ARKIN. Stochastic mechanisms in gene expression. *PNAS Proceedings of the National Academy of Sciences of the United States of America* **94**(3), 814–819 (1997).
- [229] D. MCKENZIE, P. MCGORRY, C. WALLACE, L. H. LOW, D. COPOLOV Y B. SINGH. Constructing a minimal diagnostic decision tree. *Methods of Information in Medicine* **32**(2), 161–166 (1993).
- [230] C. MEEK. Causal inference and causal explanation with background knowledge. En “Proceedings of the 11th Annual Conference on Uncertainty in Artificial Intelligence (UAI-95)”, págs. 403–410. Morgan Kaufmann Publishers (1995).

- [231] M. MEILA Y D. HECKERMAN. An experimental comparison of several clustering and initialization methods. En G. F. COOPER Y S. MORAL, editores, “Proceedings of the 14th Annual Conference on Uncertainty in Artificial Intelligence (UAI-98)”, págs. 386–395. Morgan Kaufmann (1998).
- [232] H. W. MEWES, D. FRISHMAN, U. GÜLDENER, G. MANNHAUPT, K. F. X. MAYER, M. MOKREJS, B. MORGENSTERN, M. MÜNSTERKÖTTER, S. RUDD Y B. WEIL. MIPS: a database for genomes and protein sequences. *Nucleic Acids Research* **30**(1), 31–34 (2002).
- [233] T. MICHOEL, S. MAERE, E. BONNET, A. JOSHI, Y. SAEYS, T. VAN DEN BULCKE, K. VAN LEEMPUT, P. VAN REMORTEL, M. KUIPER, K. MARCHAL Y Y. VAN DE PEER. Validating module network learning algorithms using simulated data. *BMC Bioinformatics* **8**(2) (2007).
- [234] J. MINGERS. Expert systems - rule induction with statistical data. *Journal of the Operational Research Society* págs. 39–47 (1987).
- [235] S. MIYANO, R. YAMAGUCHI, Y. TAMADA, MASAO NAGASAKI Y S. IMOTO. Gene networks viewed through two models. En S. RAJASEKARAN, editor, “Bioinformatics and Computational Biology, First International Conference, BICoB 2009. Proceedings”, vol. 5462 de “Lecture Notes in Computer Science”, págs. 54–66. Springer (2009).
- [236] E. MOLER, M. CHOW Y I. MIAN. Analysis of molecular profile data using generative and discriminative methods. *Physiological Genomics* **4**(2), 109–126 (2000).
- [237] M. MRAMOR, G. LEBAN, J. DEMSAR Y B. ZUPAN. Visualization-based cancer microarray data classification analysis. *Bioinformatics* (2007).
- [238] P. MUNTEANU Y D. CAU. Efficient score-based learning of equivalence classes of Bayesian networks. En D. A. ZIGHED, H. J. KOMOROWSKI Y J. M. ZYTKOW, editores, “Principles of Data Mining and Knowledge Discovery, 4th European Conference, PKDD 2000, Lyon, France, September 13-16, 2000, Proceedings”, vol. 1910 de “Lecture Notes in Computer Science”, págs. 96–105. Springer (2000).

- [239] K. P. MURPHY. “Dynamic Bayesian networks: Representation, inference and learning”. Tesis Doctoral, UC Berkeley, Computer Science Division (2002).
- [240] K. P. MURPHY Y S. MIAN. Modelling gene expression data using dynamic Bayesian networks. Informe técnico, Computer Science Division, University of California, Berkeley, CA (1999).
- [241] J. W. MYERS, K. B. LASKEY Y T. S. LEVITT. Learning Bayesian networks from incomplete data with stochastic search algorithms. En K. B. LASKEY Y H. PRADE, editores, “UAI ’99: Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence”, págs. 476–485. Morgan Kaufmann (1999).
- [242] N. NARIAI, S. KIM, S. IMOTO Y S. MIYANO. Using protein-protein interactions for refining gene networks estimated from microarray data by Bayesian networks. *Pacific Symposium on Biocomputing* págs. 336–47 (2004).
- [243] N. NARIAI, Y. TAMADA, S. IMOTO Y S. MIYANO. Estimating gene regulatory networks and protein–protein interactions of *Saccharomyces cerevisiae* from multiple genome-wide data. *Bioinformatics* **21**(2), 206–212 (2005).
- [244] T. NIBLETT Y I. BRATKO. Learning decision rules in noisy domains. En “Proceedings of Expert Systems ’86, The 6Th Annual Technical Conference on Research and development in expert systems III”, págs. 25–34. Cambridge University Press (1986).
- [245] W. NICHOLAS, M. F. USAMA Y S. DJORGOVSKI. Initial galaxy counts from digitized poss-II. *Astronomical Journal* **109**(6), 2401 (1995).
- [246] J. D. NIELSEN, T. KOCKA Y J. M. PEÑA. On local optima in learning Bayesian networks. En C. MEEK Y U. KJÆRULFF, editores, “UAI ’03, Proceedings of the 19th Conference in Uncertainty in Artificial Intelligence”, págs. 435–442. Morgan Kaufmann (2003).
- [247] J. NIEMI. Accuracy of the Bayesian network algorithms for inferring gene regulatory networks. Independent research projects in applied mathematics Mat-2.108, Helsinki University of Technology (2007).

- [248] N. NOVERSHTERN, Z. ITZHAKI, O. MANOR, N. FRIEDMAN Y N. KAMINSKI. A functional and regulatory map of asthma. *American Journal of Respiratory Cell and Molecular Biology*. **38**(3), 324–336 (2008).
- [249] K. NUMATA, S. IMOTO Y S. MIYANO. A structure learning algorithm for inference of gene networks from microarray gene expression data using Bayesian networks. En “Proceedings of the 7th IEEE International Conference on Bioinformatics and Bioengineering”, págs. 1280–1284. IEEE (2007).
- [250] K. OLESEN. Causal probabilistic networks with both discrete and continuous variables. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **15**(3), 275–279 (1993).
- [251] I. M. ONG, J. D. GLASNER Y D. PAGE. Modelling regulatory pathways in *E. coli* from time series expression profiles. *Bioinformatics* **18**(1), S241–8 (2002).
- [252] I. M. ONG Y D. PAGE. Inferring regulatory pathways in *E. coli* using dynamic Bayesian networks. Informe técnico 1426, University of Wisconsin-Madison (2001).
- [253] I. M. ONG Y D. PAGE. Learnability of dynamic Bayesian networks from time series microarray data. Informe técnico 1514, University of Wisconsin-Madison (2004).
- [254] A. OSAREH Y B. SHADGAR. Classification and diagnostic prediction of cancers using gene microarray data analysis. *Journal of Applied Sciences* **9**(3), 459–468 (2009).
- [255] S. OTT, A. HANSEN, S. Y. KIM Y S. MIYANO. Superiority of network motifs over optimal networks and an application to the revelation of gene network evolution. *Bioinformatics* **21**(2), 227–238 (2005).
- [256] S. OTT, S. IMOTO Y S. MIYANO. Finding optimal models for small gene networks. En R. B. ALTMAN, A. K. DUNKER, L. HUNTER, T. A. JUNG Y T. E. KLEIN, editores, “Pacific Symposium on Bio-computing”, págs. 557–567. World Scientific (2004).

- [257] S. OTT Y S. MIYANO. Finding optimal gene networks using biological constraints. En “Proceedings of the International Conference on Genome Informatics”, vol. 14, págs. 124–133 (2003).
- [258] R. PALACIOS, J. GONI, I. MARTINEZ-FORERO, J. IRANZO, J. SEPULCRE, I. MELERO Y P. VILLOSLADA. A network analysis of the human T-cell activation gene network identifies jagged1 as a therapeutic target for autoimmune diseases. *Public Library of Science (PLoS) ONE* **2**(11) (2007).
- [259] H. S. PARK, S. H. YOO Y S. B. CHO. Forward selection method with regression analysis for optimal gene selection in cancer classification. *International Journal of Computer Mathematics* **84**(5), 653–667 (2007).
- [260] M. J. PAZZANI. Searching for dependencies in Bayesian classifiers. En “Artificial Intelligence and Statistics IV”, Lecture Notes in Statistics. Springer-Verlag (1997).
- [261] J. PEARL. Fusion, propagation and structuring in belief networks. *Artificial Intelligence* **29**, 241–288 (1986).
- [262] J. PEARL. Distributed revision of composite beliefs. *Artificial Intelligence* **33**, 173–215 (1987).
- [263] J. PEARL. “Probabilistic reasoning in intelligent systems: Networks of plausible inference”. Morgan Kaufmann Publishers Inc. (1988).
- [264] E. PEELING Y A. TUCKER. Consensus gene regulatory networks: combining multiple microarray gene expression datasets. En “COMPLIFE 2007: The Third International Symposium on Computational Life Science. AIP Conference Proceedings,”, vol. 940, págs. 38–49 (2007).
- [265] D. PE’ER, A. REGEV, G. ELIDAN Y N. FRIEDMAN. Inferring sub-networks from perturbed expression profiles. *Bioinformatics* **17**(1), 215–224 (2001).
- [266] M. PELIKAN, D. E. GOLDBERG Y K. SASTRY. Bayesian optimization algorithm, decision graphs, and Occam’s razor. Illegal report no. 2000020, Illinois Genetic Algorithms Laboratory (2000).

- [267] J. M. PEÑA. Learning and validating Bayesian network models of genetic regulatory networks. En P. LUCAS, editor, “Proceedings of the 2nd European Workshop on Probabilistic Graphical Models(PGM’04)”, págs. 161–168 (2004).
- [268] J. M. PEÑA, J. BJÖRKEGREN Y J. TEGNÉR. Growing Bayesian network models of gene networks from seed genes. *Bioinformatics* **21**(2), 224–229 (2005).
- [269] J. M. PEÑA, J. BJÖRKEGREN Y J. TEGNÉR. Learning and validating Bayesian network models of gene networks. En “Advances in Probabilistic Graphical Models”, págs. 359–375. Springer-Verlag (2007).
- [270] J. M. PEÑA, J. A. LOZANO Y P. LARRAÑAGA. Learning recursive Bayesian multinets for data clustering by means of constructive induction. *Machine Learning*, 47 págs. 63–89 (2002).
- [271] J. M. PEÑA, J. A. LOZANO Y P. LARRAÑAGA. Unsupervised learning of Bayesian networks via estimation of distribution algorithms: An application to gene expression data clustering. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* **12**, 63–82 (2004).
- [272] M. A. PEOT. Geometric implications of the naïve bayes assumption. *Uncertainty in Artificial Intelligence. Proceedings of the Twelfth conference*. págs. 414–419 (1996).
- [273] F. PERNKOPF Y P. O’LEARY. Floating search algorithm for structure learning of bayesian network classifiers. *Pattern Recognition Letters* **24**(15), 2839–2848 (2003).
- [274] B. E. PERRIN, L. RALAIVOLA, A. MAZURIE, S. BOTTANI, J. MALLLET Y F. D’ALCHÉ BUC. Gene networks inference using dynamic Bayesian networks. *Bioinformatics* **19**(2), 138–148 (2003).
- [275] S. L. POMEROY, P. TAMAYO, M. GAASENBEEK, L. M. STURLA, M. ANGELO, M. E. MCCLAUGHLIN, J. Y. H. KIM, L. C. GOUMNEROVA, P. M. BLACK, C. LAU, J. C. ALLEN, D. ZAGZAG, J. M. OLSON, T. CURRAN, C. WETMORE Y J. BIEGEL. Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature* **415**, 436 – 42 (2002).

- [276] F. PROVOST Y P. DOMINGOS. Tree induction for probability-based ranking. *Machine Learning* **52**(3) (2003).
- [277] J. R. QUINLAN. Discovering rules by induction from large collection of examples. *Knowledge-base systems in the Micro Electronic Age* págs. 168–201 (1979).
- [278] J. R. QUINLAN. Simplifying decision trees. *International Journal of Man-Machine Studies* **27**(3), 221–234 (1987).
- [279] J. R. QUINLAN. “C4.5: Programs for machine learning.” Morgan Kaufmann Publishers Inc. (1993).
- [280] R DEVELOPMENT CORE TEAM. “R: A Language and Environment for Statistical Computing”. R Foundation for Statistical Computing, Vienna, Austria (2009). ISBN 3-900051-07-0.
- [281] S. RAMASWAMY, P. TAMAYO, R. RIFKIN, S. MUKHERJEE, C. H. YEANG, M. ANGELO, C. LADD, M. REICH, E. LATULIPPE, J. P. MESIROV, T. POGGIO, W. GERALD, M. LODA, E. S. LANDER Y T. R. GOLUB. Multiclass cancer diagnosis using tumor gene expression signatures. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 15149 – 54 (2001).
- [282] M. RAMONI Y P. SEBASTIANI. Learning Bayesian networks from incomplete databases. En “Proceedings of the 13th Annual Conference on Uncertainty in Artificial Intelligence (UAI-97)”, págs. 401–408. Morgan Kaufmann Publishers (1997).
- [283] Y. RAO, Y. LEE, D. JARJOURA, A. RUPPERT, C. GONG LIU, J. HSU Y J. HAGAN. A comparison of normalization techniques for microRNA microarray data. *Statistical Applications in Genetics and Molecular Biology* **7**(1), 22 (2008).
- [284] E. RAVASZ, A. L. SOMERA, D. A. MONGRU, Z. OLTVAI Y A. L. BARABÁSI. Hierarchical organization of modularity in metabolic networks. *Science* **297**(5586), 1551–5 (2002).
- [285] S. RAYCHAUDHURI, J. M. STUART Y R. B. ALTMAN. Principal components analysis to summarize microarray experiments: application

- to sporulation time series. *Pacific Symposium on Biocomputing* págs. 455–66 (2000).
- [286] G. REBANE Y J. PEARL. The recovery of causal poly-trees from statistical data. En “UAI ’87: Proceedings of the Third Annual Conference on Uncertainty in Artificial Intelligence”, págs. 175–182. Elsevier (1987).
- [287] B. REN, F. ROBERT, J. J. WYRICK, O. APARICIO, E. G. JENNINGS, I. SIMON, J. ZEITLINGER, J. SCHREIBER, N. HANNETT, E. KANIN, T. L. VOLKERT, C. J. WILSON, S. P. BELL Y R. A. YOUNG. Genome-wide location and function of DNA binding proteins. *Science* **290**(5500), 2306–9 (2000).
- [288] J. RISSANEN. Modelling by the shortest data description. *Automatica* **14**, 465–471 (1978).
- [289] J. RISSANEN. Stochastic complexity and modeling. *Annals of Statistics* **14**, 1080–1100 (1986).
- [290] R. W. ROBINSON. Counting unlabeled acyclic digraphs. En “Combinatorial mathematics V: Proceedings of the Fifth Australian Conference”, vol. 622 de “Lecture notes in mathematics”, págs. 28–43 (1977).
- [291] M. RONEN, R. ROSENBERG, B. I. SHRAIMAN Y U. ALON. Assigning numbers to the arrows: Parameterizing a gene regulation network by using accurate expression kinetics. *Proceedings of the National Academy of Sciences of the United States of America* **99**(16), 10555–60 (2002).
- [292] J. ROSAMOND Y A. ALLSOP. Harnessing the power of the genome in the search for new antibiotics. *Science* **287**, 1973–1976 (2000).
- [293] D. T. ROSS, U. SCHERF, M. B. EISEN, C. M. PEROU, C. REES, P. SPELLMAN, V. IYER, S. S. JEFFREY, M. V. DE RIJN, M. WALTHAM, A. PERGAMENSHIKOV, J. C. LEE, D. LASHKARI, D. SHALON, T. G. MYERS, J. WEINSTEIN, D. BOTSTEIN Y P. O. BROWN. Systematic variation in gene expression patterns in human cancer cell lines. *Nature Genetics* **24**(3), 227–35 (2000).

- [294] R. RUMÍ, A. SALMERÓN Y S. MORAL. Estimating mixtures of truncated exponentials in hybrid Bayesian networks. *TEST: An Official Journal of the Spanish Society of Statistics and Operations Research* **15**(2), 397–421 (2006).
- [295] K. SACHS, O. PEREZ, D. PE’ER, D. A. LAUFFENBURGER Y G. P. NOLAN. Causal protein-signaling networks derived from multiparameter single-cell data. *Science* **308**(5721), 523–529 (2005).
- [296] M. SAHAMI. Learning limited dependence Bayesian classifiers. *Second International Conference on Knowledge Discovery in Databases* págs. 335–338 (1996).
- [297] A. SALMERÓN. Algoritmos de propagación II: Métodos de Monte Carlo. En J. A. GÁMEZ Y J. M. PUERTA, editores, “Sistemas Expertos Probabilísticos”, Ciencia y Técnica num. 20, págs. 65–88. Universidad de Castilla-La Mancha, 1 ed. (1998).
- [298] S. SALZBERG, R. CHANDAR, H. FORF, S. MURTH Y R. WHITE. Decision trees for automated identification of cosmic-ray hits in hubble space telescope images. *Publications of the Astronomical Society of the Pacific* **107**, 1–10 (1995).
- [299] C. J. SAVOIE, S. ABURATANI, S. WATANABE, Y. EGUCHI, S. MUTA, S. IMOTO, S. MIYANO, S. KUHARA Y K. TASHIRO. Use of gene networks from full genome microarray libraries to identify functionally relevant drug-affected genes and gene regulation cascades. *DNA Research* **10**(1), 19–25 (2003).
- [300] M. SCHENA, D. SHALON, R. W. DAVIS Y P. O. BROWN. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**(5235), 467–70 (1995).
- [301] U. SCHERF, D. T. ROSS, M. WALTHAM, L. H. SMITH, J. K. LEE, L. TANABE, K. W. KOHN, W. C. REINHOLD, T. G. MYERS, D. T. ANDREWS, D. A. S. AND MICHAEL B. EISEN, E. A. SAUSVILLE, Y. POMMIER, D. BOTSTEIN, P. O. BROWN, Y J. WEINSTEIN. A gene expression database for the molecular pharmacology of cancer. *Nature Genetics* **24**(3), 236–244 (2000).

- [302] G. SCHWARZ. Estimating the dimension of a model. *Annals of Statistics* **6**, 461–464 (1978).
- [303] E. SEGAL, D. PE’ER, A. REGEV, D. KOLLER Y N. FRIEDMAN. Learning module networks. En “Proceedings Nineteenth Conference on Uncertainty in Artificial Intelligence (UAI)”, págs. 525–534 (2003).
- [304] E. SEGAL, D. PE’ER, A. REGEV, D. KOLLER Y N. FRIEDMAN. Learning module networks. *Journal of Machine Learning Research* **6**, 557–588 (2005).
- [305] E. SEGAL, M. SHAPIRA, A. REGEV, D. PE’ER, D. BOTSTEIN, D. KOLLER Y N. FRIEDMAN. Module networks: identifying regulatory modules and their condition-specific regulators from gene expression data. *Nature Genetics* **34**(2), 166–76 (2003).
- [306] E. SEGAL, H. WANG Y D. KOLLER. Discovering molecular pathways from protein interaction and gene expression data. *Bioinformatics* **19**(1) (2003).
- [307] I. SEN, M. P. VERDICCHIO, S. JUNG, R. TREVINO, M. BITTNER Y S. KIM. Context-specific gene regulations in cancer gene expression data. En “Pacific Symposium on Biocomputing”, vol. 14, págs. 75–86 (2009).
- [308] S. SHAH Y A. KUSIAK. Cancer gene search with data-mining and genetic algorithms. *Computers in Biology and Medicine* **37**(2), 251–261 (2007).
- [309] C. E. SHANNON. A mathematical theory of communication. *Bell System Technical Journal* **27**, 379–423 (1948).
- [310] P. P. SHENOY Y G. SHAFER. Axioms for probability and belief-function propagation. En “Uncertainty in Artificial Intelligence”, págs. 169–198. Morgan Kaufmann (1990).
- [311] P. P. SHENOY Y G. SHAFER. Probability propagation. En “Annals of Mathematics and Artificial Intelligence”, vol. 2, págs. 327–351 (1990).
- [312] S. SHIMONY Y E. CHARNIAK. A new algorithm for finding MAP assignments to belief network. En “Proceedings of the 6th Annual

- Conference on Uncertainty in Artificial Intelligence (UAI-91)", págs. 185–196. Elsevier Science (1991).
- [313] S. SHIMOZONO, A. SHINOHARA, T. SHINOHARA, S. MIYANO, S. KUHARA Y S. ARIKAWA. Knowledge acquisition from amino acid sequences by machine learning system BONSAI. *Transactions on Information Processing Society of Japan* **35**(10), 2009–2018 (1994).
 - [314] M. A. SHIPP, K. ROSS, P. TAMAYO, A. P. WENG, J. L. KUTOK, R. C. T. AGUIAR, M. GAASENBEEK, M. ANGELO, M. REICH, G. S. PINKUS, T. S. RAY, M. A. KOVAL, K. W. LAST, A. NORTON, T. A. LISTER Y J. MESIROV. Diffuse large B-cell lymphoma outcome prediction by gene-expression profiling and supervised machine learning. *Nature Medicine* **8**, 68 – 74 (2002).
 - [315] B. SIERRA Y P. LARRAÑAGA. Predicting survival in malignant skin melanoma using Bayesian networks automatically induced by genetic algorithms. An empirical comparison between different approaches. *Artificial Intelligence in Medicine* **14**(1-2), 215–230 (1998).
 - [316] D. SINGH, P. G. FEBBO, K. ROSS, D. G. JACKSON, J. MANOLA, C. LADD, P. TAMAYO, A. A. RENSHAW, A. V. D'AMICO, J. P. RICHIE, E. S. LANDER, M. LODA, P. W. KANTOFF, T. R. GOLUB Y W. R. SELLERS. Gene expression correlates of clinical prostate cancer behavior. *Cancer Cell* **1**, 203 – 9 (2002).
 - [317] M. SINGH Y M. VALTORTA. Construction of Bayesian network structures from data: A brief survey and an efficient algorithm. *International Journal of Approximate Reasoning* **12**(2), 111–131 (1995).
 - [318] D. K. SLONIM, P. TAMAYO, J. P. MESIROV, T. R. GOLUB Y E. S. LANDER. Class prediction and discovery using gene expression data. En "Proceedings of the fourth annual international conference on Computational molecular biology (RECOMB 2000)", págs. 263–272 (2000).
 - [319] P. T. SPELLMAN, G. SHERLOCK, M. Q. ZHANG, V. R. IYER, K. ANDERS, M. B. EISEN, P. O. BROWN, D. BOTSTEIN Y B. FUTCHER. Comprehensive identification of cell cycle-regulated genes of the

- yeast *Saccharomyces cerevisiae* by microarray hybridization. *Molecular Biology of the Cell* **9**, 3273–3297 (1998).
- [320] P. SPIRTEs Y C. GLYMOUR. An algorithm for the fast recovery of sparse casual graphs. *Social Science Computer Review* **9**(1), 62–72 (1991).
- [321] P. SPIRTEs, C. GLYMOUR Y R. SCHEINES. “Causation, prediction, and search”. Springer-Verlag (1993).
- [322] P. SPIRTEs, C. GLYMOUR, R. SCHEINES, S. KAUFFMAN, V. AIMALE Y F. WIMBERLY. Constructing Bayesian network models of gene expression networks from microarray data. En “Proceedings of the Atlantic Symposium on Computational Biology, Genome Information Systems and Technology” (2000).
- [323] P. SPIRTEs Y C. MEEK. Learning Bayesian networks with discrete variables from data. En “Proceedings of First International Conference on Knowledge Discovery and Data Mining (KDD)”, págs. 294–299. Morgan Kaufmann (1995).
- [324] W. STACKLIES, H. REDESTIG, M. SCHOLZ, D. WALTHER Y J. SELBIG. pcaMethods—a bioconductor package providing PCA methods for incomplete data. *Bioinformatics* **23**(9), 1164–1167 (2007).
- [325] H. STECK. On the use of skeletons when learning in Bayesian networks. En “Proceedings of the 16th Annual Conference on Uncertainty in Artificial Intelligence (UAI-00)”, págs. 558–565. Morgan Kaufmann Publishers (2000).
- [326] M. STONE. Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society B* **36**(1), 111–147 (1974).
- [327] M. P. STYCZYNSKI Y G. STEPHANOPOULOS. Overview of computational methods for the inference of gene regulatory networks. *Computers & Chemical Engineering* **29**(3), 519–534 (2005).
- [328] J. SUZUKI. A construction of Bayesian networks from databases based on the MDL principle. En “Proceedings of the 9th Annual Conference on Uncertainty in Artificial Intelligence”, págs. 266–273. Morgan Kaufmann Publishers (1993).

- [329] J. SUZUKI. Learning Bayesian belief networks based on the minimum description length principle: An efficient algorithm using the B & B technique. En “International Conference on Machine Learning”, págs. 462–470 (1996).
- [330] J. A. SWETS Y R. M. PICKETT. Evaluation of diagnostic systems: Methods from signal detection theory. *Medical Physics* **10**(2), 266–267 (1983).
- [331] Y. TAMADA, H. BANNAI, S. IMOTO, T. KATAYAMA, M. KANEHISA Y S. MIYANO. Estimating gene networks from gene expression data and evolutionary information using Bayesian network models. En “Proceedings 15th International Conference on Genome Informatics, Posters and Software Demonstrations (GIW’04)” (2004).
- [332] Y. TAMADA, S. IMOTO, K. TASHIRO, S. KUHARA Y S. MIYANO. Identifying drug active pathways from gene networks estimated by gene expression data. *Genome Informatics* **16**(1), 182–191 (2005).
- [333] Y. TAMADA, S. KIM, H. BANNAI, S. IMOTO, K. TASHIRO, S. KUHARA Y S. MIYANO. Combining gene expression data with DNA sequence information for estimating gene networks using Bayesian network model. En “Proceedings of the International Conference on Genome Informatics”, vol. 14, págs. 352–353 (2003).
- [334] Y. TAMADA, S. KIM, H. BANNAI, S. IMOTO, K. TASHIRO, S. KUHARA Y S. MIYANO. Estimating gene networks from gene expression data by combining Bayesian network model with promoter element detection. *Bioinformatics* **19**(2), 227–236 (2003).
- [335] B. THIESSON, C. MEEK, D. M. CHICKERIN Y D. HECKERMAN. Learning mixtures of DAG models. En G. F. COOPER Y S. MORAL, editores, “Proceeding of the 14th Conference on Uncertainty in Artificial Intelligence (UAI-98)”, págs. 504–513. Morgan Kaufmann, San Francisco, CA (1998).
- [336] J. TIAN. A branch-and-bound algorithm for MDL learning Bayesian networks. En C. BOUTILIER Y M. GOLDSZMIDT, editores, “Proceedings of the 16th Annual Conference on Uncertainty in Artificial Intelligence (UAI-00)”, págs. 580–588. Morgan Kaufmann Publishers (2000).

- [337] I. M. TIENDA-LUNA, Y. HUANG, Y. YIN, DIEGO Y CARMEN. Uncovering gene regulatory networks from time-series microarray data with variational Bayesian structural expectation maximization. *EURASIP Journal on Bioinformatics and Systems Biology* **2007** (2007).
- [338] I. M. TIENDA-LUNA, Y. YIN, M. C. CARRION, Y. HUANG, H. CAI, M. SANCHEZ Y Y. WANG. Inferring the skeleton cell cycle regulatory network of malaria parasite using comparative genomic and variational Bayesian approaches. *Genetica* **132**(2), 131–142 (2007).
- [339] N. TISHBY, F. C. PEREIRA Y W. BIALEK. The information bottleneck method. En “The 37th annual Allerton Conference on Communication, Control, and Computing”, págs. 368–377 (1999).
- [340] J. B. TOBLER, M.N.MOLLA, E. F. NUWAYSIR, R. D. GREEN Y J. W. SHAVLIK. Evaluating machine learning approaches for aiding probe selection for gene-expression arrays. En “Proceedings of the Tenth International Conference on Intelligent Systems for Molecular Biology ISMB (Supplement of Bioinformatics)”, págs. 164–171 (2002).
- [341] O. TROYANSKAYA, M. CANTOR, G. SHERLOCK, P. BROWN, T. HASTIE, R. TIBSHIRANI, D. BOTSTEIN Y R. ALTMAN. Missing value estimation methods for DNA microarrays. *Bioinformatics* **17**(6), 520–525 (2001).
- [342] A. TUCKER, V. VINCIOTTI, P. A. C. ’T HOEN Y X. LIU. Bayesian network classifiers for time-series microarray data. En A. F. FAMILI, J. KOK, J. M. PEÑA, A. SIEBES Y A. J. FEELDERS, editores, “Advances in Intelligent Data Analysis VI, 6th International Symposium on Intelligent Data Analysis, IDA 2005, Proceedings”, vol. 3646 de “Lecture Notes in Computer Science”, págs. 475–485. Springer (2005).
- [343] N. TURTON, D. JUDAH, J. RILEY, R. DAVIES, D. LIPSON, J. STYLES, A. SMITH Y T. GANT. Gene expression and amplification in breast carcinoma cells with intrinsic and acquired doxorubicin resistance. *Oncogene* **20**(11), 1300–1306 (2001).
- [344] L. J. VAN ’T VEER, H. DAI, M. J. VAN DE VIJVER, Y. D. HE, A. A. HART, M. MAO, H. L. PETERSE, K. VAN DER KOOY, M. J.

- MARTON, A. T. WITTEVEEN, G. J. SCHREIBER, R. M. KERKHOVEN, C. ROBERTS, P. S. LINSLEY, R. BERNARDS Y S. H. FRIEND. Gene expression profiling predicts clinical outcome of breast cancer. *Nature* **415**(6871), 530–536 (2002).
- [345] T. VERMA Y J. PEARL. Causal networks: Semantics and expressiveness. En R. D. SHACHTER, T. S. LEVITT, L. N. KANAL Y J. F. LEMMER, editores, “Uncertainty in Artificial Intelligence 4”, págs. 69–76. North-Holland (1990).
- [346] T. VERMA Y J. PEARL. Equivalence and synthesis of causal models. En “Proceedings of the 6th Annual Conference on Uncertainty in Artificial Intelligence (UAI-91)”, págs. 220–227. Elsevier Science (1991).
- [347] P. WALLEY. Inferences from multinomial data; learning about a bag of marbles (with discussion). *Journal of the Royal Statistical Society, Series B* **58**, 3–57 (1996).
- [348] F. WANG, D. PAN Y J. DING. A new approach combined fuzzy clustering and Bayesian networks for modeling gene regulatory networks. En “BMEI ’08: Proceedings of the 2008 International Conference on Bio-Medical Engineering and Informatics”, págs. 29–33. IEEE Computer Society (2008).
- [349] M. WANG, Z. CHEN Y S. CLOUTIER. A hybrid Bayesian network learning method for constructing gene networks. *Computational Biology and Chemistry* **31**(5-6), 361–372 (2007).
- [350] S. M. WEISS Y C. A. KULIKOWSKI. “Computer systems that learn: Classification and prediction. Methods from statistics, neural nets, machine learning and expert systems”, cap. Chapter 2: How to estimate the True Performance of a Learning System, págs. 17–49. Morgan Kaufmann Publishers Inc. (1991).
- [351] A. V. WERHLI, M. GRZEGORCZYK Y D. HUSMEIER. Comparative evaluation of reverse engineering gene regulatory networks with relevance networks, graphical Gaussian models and Bayesian networks. *Bioinformatics* (2006).
- [352] A. V. WERHLI Y D. HUSMEIER. Reconstructing gene regulatory networks with Bayesian networks by combining expression data with