บรรณานุกรม

ไก้รุ่ง เฮงพระพรหม. (2553). *การเลือกคุณลักษณะสำหรับการพยากรณ์ค่าที่ขาดหายสำหรับข้อมูลหลายมิติและการประยุกต์สำหรับข้อมูลไมโครอาร์เรย์*. ดุษฎีนิพนธ์ปรัชญาดุษฎีบัณฑิต, สาขาวิชาเทคโนโลยีสารสนเทศ, บัณฑิตวิทยาลัย, มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าพระนครเหนือ.

Acuna, E., & Rodriguez, C. (2004). *A meta analysis study of outlier detection methods in classification*. Department of Mathematics. University of Puerto Rico at Mayaguez. Retrieved December 7, 2012, from http*://academic.uprm.edu/~eacuna/paperout.pdf*

Alma, O. G., & Bulut, E. (2012). Genetic algorithm based variable selection for partial least squares regression using ICOMP criteria. *Asian Journal of Mathematics and Statistics*, *5*(3), 82-92.

Alon, U., Barkai, N., Notterman, D. A., Gish, K., Ybarra, S., Mack, D., & Levine, A. J. (1999). *Data pertaining to the article 'Broad patterns of gene expression revealed by clustering of tumor and normal colon tissues probed by oligonucleotide arrays'* [Data set]. Retrieved August 6, 2012, from http://genomics-pubs.princeton.edu/oncology/affydata/index.html

Alsberg, B. K., Kell, D. B., & Goodacre, R. (1998). Variable selection in discriminant partial least-squares analysis. *Anal. Chem., 70*(19), 4126-4133.

Anzanello, M., Albin, S. L., & Chaovalitwongse, W. A. (2009). Selecting the best variables for classifying production batches into two quality levels. *Chemometrics and Intelligent Laboratory Systems*, *97*, 111-117.

Arriaga, J. M., Levy, E. M., Bravo, A. I., Bayo, S. M., Amat, M., Aris, M., Hannois, A., Bruno, L., Roberti, M. P., Loria, F. S., Pairola, A., Huertas, E., Mordoh, J., Bianchini, M. (2012). Metallothionein expression in colorectal cancer: relevance of different isoforms for tumor progression and patient survival. *Human Pathology, 43*, 197-208.

Barker, M., & Rayens, W. (2003). Partial least squares for discrimination. *Journal of Chemometrics*, *17*, 166-173.

Ben-Dor, A., Bruhn, L., Friedman, N., Nachman, I., Schummer, M., & Yakhini, Z. (2000). Tissue Classification with Gene Expression Profiles, *Journal of Computational Biology*, *7*, 559-583.

Boulesteix, A.-L. (2004). PLS dimension reduction for classification with microarray data. *Statistical Applications in Genetics and Molecular Biology*, *3*(1), Article 33.

Boulesteix, A.-L. (2006). Reader's reaction to "Dimension reduction for classification with gene expression microarray data" by Dai et al (2006), *Statistical Applications in Genetics and Molecular Biology*, *5*(1), Article 16.

Boulesteix, A.-L., & Strimmer, K. (2006). Partial least squares: A versatile tool for the Analysis of high-dimensional genomic data. *Briefings in Bioinformatics*, *8*(1), 32-44.

Breikers, G., van Breda, S. G., Bouwman, F. G., van Herwijnen, M. H., Renes, J., Mariman, E. C., Kleinjans, J. C., & van Delft, J. H. (2006). for nutritional health effects on colorectal cancer in the mouse as revealed by proteomics analysis. *Proteomics, 6*, 2844-2852.

Cai, W., Li, Y., & Shao, X. G. (2008). A variable selection method based on uninformative variable elimination for multivariate calibration of near-infrared spectra. *Chemometrics and Intelligent Laboratory Systems*, *90*, 188-194.

Centner, V., & Massart, D.-L. (1996). Elimination of uninformative variables for multivariate calibration. *Analytical Chemistry*, *68*(21), 3851-3858.

Chen, Y., & Nguyen D. (n.d.). *Identification of relevant genes from microarray experiments based on partial least squares weights: Application to cancer genomics.* Retrieved June 20, 2012, from http://dnguyen.ucdavis.edu/.html/PLS_VS.pdf

Chen, Y. (2008). *Statistical approaches for detection of relevant genes and pathway in analysis of gene expression data*. Doctoral Dissertation, Statistics, University of California.

Cho, J.-H., Lee, D., Park, J. H., Kim, K., & Lee, I.-B. (2002). Optimal approach for classification of acute leukemia subtypes based on gene expression data. *Biotechnology Progress, 18*(4)*,* 847-854.

Chong, I.-G., & Jun, C.-H. (2005). Performance of some variable selection methods when multicollinearity is present. *Chemometrics and Intelligent Laboratory Systems*, *78*, 103-112.

Dai, J. J., Lieu, L. , & Rocke, D. (2006). Dimension reduction for classification with gene expression microarray data. *Statistical Applications in Genetics and Molecular Biology*, *5*(1), Article 6.

Dash, M., & Liu, H. (1997). Feature selection for classification. *Intelligent Data Analysis*, *1*, 131-156.

De Lozanne, A., & Spudich, J. A. (1987). Disruption of the Dictyostelium myosin heavy chain gene by homologous recombination [Abstract]. *Science, 236*(4805), 1086-1091.

Forman, G. (2003). An extensive empirical study of feature selection metrics for text classification. *Journal of Machine Learning Research*, *3*, 1289-1305.

Fujarewicz, K., & Wiench, M. (2003). Selecting differentially expressed genes for colon tumor classification. *International Journal of Applied Mathematics* and Computer *Science, 13*(3), 327–335.

Gordon, R. A. (2012). *Applied Statistics for the Social and Health Sciences*. New York: Routledge.

Gosselin, R., Rodrigue, D., & Duchesne, C. (2010). A bootstrap-VIP approach for selecting wavelength intervals in spectral imaging application. *Chemometrics and Intelligent Laboratory Systems*, *100*, 12-21.

Gutkin, M., Shamir, R., & Dror, G. (2009). SlimPLS: A method for feature selection in gene expression-based disease classification. *PLoS ONE*, *4*(7), e6416.

Guyon, I. (2008). Practical feature selection from correlation to causality. In F. Fogelman-Soulié, D. Perrotta, J. Piskorski, & R. Steinberger (Eds.), *Mining Massive Data Sets for Security: Advances in Data Mining, Search, Social Networks and Text Mining, and Their Applications to Security* (pp. 27–43). Amsterdam: IOS Press.

Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, *3*, 1157–1182.

Guyon, I., Weston, J., Barnhill, S., & Vapnik, V. (2002). Gene selection for cancer classification using support vector machines. *Machine Learning*, *46*(1-3), 389–422.

Han, Q.-J., Wu, H.-L., Cai, C.-B., Xu, L., & Yu, R.-Q. (2008). An ensemble of monte carlo uninformative variable elimination for wavelength selection. *Analytica Chimica Acta*, *612*, 121-125.

Hilario, M., & Kalousis, A. (2008). Approaches to dimensionality reduction in proteomic biomarker studies. *Briefings in Bioinformatics*, *9*(2), 102-118.

Höskuldsson, A. (1988). PLS regression methods. *Journal of Chemometrics*, *2*, 211-228.

Huang, X., & Pan, W. (2003). Linear regression and two-class classification with gene expression data. *Bioinformatics, 19*(16), *2072–2078.*

Janecek, A. (2009). *Efficient feature reduction and classification methods: applications in drug discovery and email categorization*. Doctoral dissertation, University of Vienna.

Ji, G., Yang, Z., & You, W. (2011). PLS-based gene selection and identification of tumor-specific genes. *IEEE Transactions on Systems Man and Cybernetics-Part C: Applications and Reviews*, *41*(6), 830-841.

Johansson, D., Lindgren, P., & Berglund, A. (2003). A multivariate approach applied to microarray data for identification of genes with cell cycle-coupled transcription. *Bioinformatics*, *19*(4), 467-473.

John, G. H., Kohavi, R., & Pfleger, K. (1994). Irrelevant features and the subset selection problem. In W. W. Cohen & H. Hirsh (Eds.), *Machine Learning: Proceedings of the Eleventh International Conference* (pp. 121-129). CA: Morgan Kaufmann Publishers.

Keely, P., Parise, L., &Juliano, R. (1998). Integrins and GTPases in tumour cell growth, motility and invasion. *Trends in Cell Biology, 8*, 101-107.

Kira, K., & Rendell, L.A. (1992) The feature selection problem: Traditional methods and a new algorithm. In *Proceedings AAAI-92*, *San Jose, CA* (pp. 129-134). MA: MIT Press)

Koller, D., & Sahami, M. (n.d.). Toward optimal feature selection. Retrieved December 6, 2012, from http://ilpubs.stanford.edu:8090/208/1/1996-77.pdf

Krzanowski, W. J., & Hand, D. J. (2009). A simple method for screening variables before clustering microarray data. *Computational Statistics and Data Analysis*, *53*, 2747-2753.

Ladha, L., & Deepa, T. (2011). Feature selection methods and algorithms. *International Journal on Computer Science and Engineering*, *3(*5), 1787-1797.

Lazraq, A., Cléroux, R., & Gauchi, J.–P. (2003). Selecting both latent and explanatory variables in the PLS1 regression model. *Chemometrics and Intelligent Laboratory Systems*, *66*, 117-126.

Li, B., Morris, J., & Martin, E. B. (2002). Model selection for partial Least squares regression. *Chemometrics and Intelligent Laboratory Systems*, *64*, 79-89.

Li, G.-Z., & Zeng, X.-Q. (2009). Feature selection for partial least square based dimension reduction. In A. Abraham et al. (Eds.), *Foundations of computational intelligence* (Vol. 5, p. 3-37). Springer-Verlag Berlin Heidelberg.

Li, G.-Z., Zhao, R.-W., Qu, H.-N., & You, M. (2012). Model selection for partial Least squares based dimension reduction. *Pattern Recognition Letters*, *33*, 524-529.

Liu, H. & Motoda, H. (1998). *Feature Selection for Knowledge Discovery and Data Mining*. MA: Kluwer Academic Publishers Norwell.

Ma, Y. L., Peng, J. Y., Liu, W. J., Zhang, P., Huang, L., Gao, B., Shen, T., Zhou, Y., Chen, H., Chu, Z., Zhang, M., & Qin, H. *(*2009). Proteomics identification of desmin as a potential oncofetal diagnostic and prognostic biomarker in colorectal cancer. *Molecular & Cellular Proteomics, 8*, 1878–1890.

Man, M. Z., Dyson, G., Johnson, K., & Liao, B. (2004). Evaluating methods for classifying expression data. *Journal of Biopharmaceutical Statistics, 14*(4), 1065–1084.

Manne, R. (1987). Analysis of two partial-least-squares algorithms for multivariate calibration. *Chemometrics and Intelligent Laboratory Systems*, *2*, 187-197.

Mehmood, T., Liland, K. H., Snipen, L., & Sæbø, S. (2012). A review of variable selection methods in partial least squares regression. *Chemometrics and Intelligent Laboratory Systems*, *118*, 62-69.

Moros, J., Kuligowski, J., Quintas,G., Garrigues, S., & Guardia, M. (2008). New cut-off criterion for uninformative variable elimination in multivariate calibration of near-infrared spectra for the determination of heroin in illicit street drugs. *Analytica Chimica Acta*, *630*, 150-160.

Narendra, P. M., & Fukunaga, K. (1977). A branch and bound algorithm for feature selection. *IEEE Transactions on Computers*, *C-26*(9), 917-922.

Nguyen, D., & Rocke, D. M. (2002a). Tumor classification by partial least squares using microarray gene expression data. *Bioinformatics*, *18*(1), 39–50.

Nguyen, D., & Rocke, D. M. (2002b). Multi-class cancer classification via partial least squares with gene expression profiles. *Bioinformatics*, *18*(9), 1216–1226.

Novaković, J., Strbac, P., & Bulatović, D. (2011). Toward optimal feature selection using ranking methods and classification algorithms. *Yugoslav Journal of Operations Research*, *21*(1), 119-135.

Pérez-Enciso, M., & Tenenhaus, M. (2003). Prediction of clinical outcome with microarray data: a partial least squares discriminant analysis (PLS-DA) approach. *Human Genetics*, *112*, 581–592.

Rajalahti, T., & Kvalheim, O. M. (2011), Multivariate data analysis in pharmaceutics: A tutorial review, *International Journal of Pharmaceutics*, *417*, 280–290.

Rosipal, R., & Krämer, N. (2006). Overview and recent advances in partial least squares. In C. Saunders et al. (Eds.), *Lecture Notes in Computer Science* (pp.34-51). Springer-Verlag Berlin Heidelberg.

Rovner, A. S., Murphy, R. A., & Owens, G. K. (1986). Expression of smooth muscle and nonmuscle myosin heavy chains in cultured vascular smooth muscle cells. *The Journal of Bilological Chemistry, 261*(31), 14740-14745.

Russolillo, G. (2009). *Partial least squares methods for non-metric data*. Doctoral Dissertation, Dipartimento di Matematica e Statistica, Universita degli Studi di Napoli Federico II. Napoli.

Saeys, Y., Inza, I., & Larrañaga, P. (2007). A review of feature selection techniques in bioinformatics. *Bioinformatics*, *23*(19), 2507–2517.

Shailubhai, K., Yu, H. H., Karunanandaa, K., Wang, J. Y., Eber, S. L., Wang, Y.,Joo, N. S., Kim, H. D., Miedema, B. W., Abbas, S. Z., Boddupalli, S. S., Currie, M. G., & Fort, L. R.(2000). Uroguanylin treatment suppresses polyp formation in the Apc(Min/+) mouse and induces apoptosis in human colon adenocarcinoma cells via cyclic GMP. *Cancer Research*, *60*(18), 5151-5157.

Shao, X. G., Wang, F., Chen, D., & Su, Q. D. (2004). A method for near-infrared spectral calibration of complex plant samples with wavelet transform and elimination of uninformative variables. *Analytical Bioanalytical Chemistry*, *378*(5), 1382-1387.

Simons, M., & Rosenberg, R. D. (1992). Antisense nonmuscle myosin heavy chain and c-myb oligonucleotides suppress smooth muscle cell proliferation in vitro. *Circulation Research, 70*(4), 835–843.

Sun, X.-M., Yu, X.-P., Liu, Y., Xu, L., & Di, D.-L. (2012). Combining bootstrap and uninformative variable elimination: Chemometric identification of metabonomic biomarkers by nonparametric analysis of discriminant partial least squares. *Chemometrics and Intelligent Laboratory Systems*, *115*, 37-43.

Tan, P.-N., Steinbach, M., & Kumar, V. (2006). *Introduction to Data Mining*. Boston: Addison Wesley.

Tan, Y., Shi, L., Tong, W., Hwang, G. T. G., & Wang, C. (2004). Multi-class tumor classification by discriminant partial least squares using microarray gene expression data and assessment of classification models. *Computational Biology and Chemistry*, *28*, 235–244.

Teófilo, R. F., Martins, J. P. A., & Ferreira, M. M. C. (2008). Sorting variables by using informative vectors as a strategy for feature selection in multivariate regression. *Journal of Chemometrics*, *23*, 32-48.

Wang, L., Zhu, J., & Zou, H. (2008). Hybrid huberized support vector machines for microarray classification and gene selection. *Bioinformatics*, *24*(3), 412–419.

Wold, S., Sjöström, M., & Eriksson, L. (2001). PLS-regression: A basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems*, *58*, 109-130.

Yam, J. W., Chan, K. W., & Hsiao, W. L. (2001). Suppression of the tumorigenicity of mutant p53-transformed rat embryo fibroblasts through expression of a newly cloned rat nonmuscle myosin heavy chain-B, *Oncogene*, *20*(1), 58-68.

Yang, A. J., & Song, X. Y. (2010). Bayesian variable selection for disease classification using gene expression data. *Bioinformatics*, *26*(2), 215–222.

Yeniay, Ö., & Göktaş, A. (2002). A comparison of partial least squares regression with other prediction methods. *Hacettepe Journal of Mathematics and Statistics*, *31*, 99-111.

Yu, L., & Liu, H. (2003). Feature selection for high-dimensional data: A fast correlation-based filter solution. In *Proceedings of the Twentieth International Conference on Machine Learning (ICML-2003)*, Washington DC (pp. 856-863). n.p.

Yu, L., & Liu, H. (2004). Efficient feature selection via analysis of relevance and redundancy. *Journal of Machine Learning Research*, *5*,1205–1224.

Zeng, X. Q., Li, G. Z., & Wu, G. F. (n.d.). On the number of partial least squares components in dimension reduction for tumor classification. Retrieved September 28, 2012, from http://levis.tongji.edu.cn/gzli/publication/biodm07.pdf

Zeng, X. Q., Li, G. Z., Wu, G. F., Yang, J. Y., & Yang, M. Q. (2008). Irrelevant gene elimination for partial least squares based dimension reduction by using feature probes. *International Journal of Data Mining & Bioinformatics*, (in press).

Zheng, H., & Zhang, Y. (2008). Feature selection for high-dimensional data in astronomy. *Advances in Space Research*, *41*, 1960-1964.