INTEGRATED SMOOTHED LOCATION MODEL AND DATA REDUCTION APPROACHES FOR MULTI VARIABLES CLASSIFICATION

HASHIBAH BINTI HAMID

DOCTOR OF PHILOSOPHY
UNIVERSITI UTARA MALAYSIA
2014

Permission to Use

In presenting this thesis in fulfilment of the requirements for a postgraduate degree from Universiti Utara Malaysia, I agree that the Universiti Library may make it freely available for inspection. I further agree that permission for the copying of this thesis in any manner, in whole or in part, for scholarly purpose may be granted by my supervisor(s) or, in their absence, by the Dean of Awang Had Salleh Graduate School of Arts and Sciences. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to Universiti Utara Malaysia for any scholarly use which may be made of any material from my thesis.

Requests for permission to copy or to make other use of materials in this thesis, in whole or in part, should be addressed to:

Dean of Awang Had Salleh Graduate School of Arts and Sciences
UUM College of Arts and Sciences
Universiti Utara Malaysia
06010 UUM Sintok

Abstrak

Model Lokasi Terlicin merupakan satu peraturan klasifikasi yang berurusan dengan campuran pembolehubah selanjar dan pembolehubah binari secara serentak. Peraturan ini mendiskriminasi kumpulan dalam bentuk berparameter menggunakan taburan bersyarat pembolehubah selanjar dengan berpandukan kepada setiap corak pembolehubah binari. Bagi menjalankan analisis pengklasifikasian yang praktikal, objek mesti terlebih dahulu disusun ke dalam sel jadual multinomial yang dijana daripada pembolehubah binari. Seterusnya, parameter dalam setiap sel akan dianggarkan dengan menggunakan objek-objek yang disusun. Walau bagaimanapun, dalam kebanyakan situasi, parameter yang dianggarkan adalah lemah jika bilangan binari adalah besar berbanding dengan saiz sampel. Pembolehubah binari yang besar akan mencipta terlalu banyak sel-sel multinomial yang kosong, yang membawa kepada masalah sparsiti tinggi dan akhirnya memberikan prestasi yang terlampau buruk untuk peraturan yang telah dibangunkan. Pada senario paling teruk, peraturan tidak dapat dibangunkan. Bagi mengatasi kekurangan tersebut, kajian ini mencadangkan strategi baru untuk mengekstrak pembolehubah yang mencukupi bagi menyumbang kepada prestasi peraturan yang optimum. Gabungan dua teknik pengekstrakan diperkenalkan, iaitu 2PCA dan PCA+MCA dengan titik aras baru nilai eigen dan jumlah varians terterang, untuk menentukan kecukupan pembolehubah terekstrak yang mendorong kepada kadar salah klasifikasi yang minimum. Hasil daripada teknik-teknik pengekstrakan ini digunakan untuk membangunkan model lokasi terlicin, yang kemudiannya menghasilkan dua pendekatan klasifikasi baharu iaitu 2PCALM dan 2DLM. Bukti berangka daripada kajian simulasi menunjukkan bahawa kadar salah klasifikasi yang dihitung menunjukkan tiada perbezaan yang signifikan di antara teknik-teknik pengekstrakan dalam data normal dan tidak normal. Walau bagaimanapun, kedua-dua pendekatan yang dicadangkan adalah sedikit terjejas bagi data tidak normal dan terjejas teruk bagi kumpulan yang amat bertindan. Kajian terhadap beberapa set data sebenar menunjukkan bahawa kedua-dua pendekatan ini adalah kompetitif, dan lebih baik daripada kaedah klasifikasi lain yang sedia ada. Keseluruhan penemuan mendedahkan bahawa kedua-dua pendekatan yang dicadangkan dapat dipertimbangkan sebagai penambahbaikan terhadap model lokasi, dan alternatif kepada kaedah klasifikasi lain terutamanya dalam mengendalikan pembolehubah campuran dengan saiz binari yang besar.

Kata kunci: Peraturan Klasifikasi, Model Lokasi Terlicin, Analisis Komponen Utama, Analisis Kesepadanan Berganda, Pembolehubah Campuran.

Abstract

Smoothed Location Model is a classification rule that deals with mixture of continuous variables and binary variables simultaneously. This rule discriminates groups in a parametric form using conditional distribution of the continuous variables given each pattern of the binary variables. To conduct a practical classification analysis, the objects must first be sorted into the cells of a multinomial table generated from the binary variables. Then, the parameters in each cell will be estimated using the sorted objects. However, in many situations, the estimated parameters are poor if the number of binary is large relative to the size of sample. Large binary variables will create too many multinomial cells which are empty, leading to high sparsity problem and finally give exceedingly poor performance for the constructed rule. In the worst case scenario, the rule cannot be constructed. To overcome such shortcomings, this study proposes new strategies to extract adequate variables that contribute to optimum performance of the rule. Combinations of two extraction techniques are introduced, namely 2PCA and PCA+MCA with new cut-points of eigenvalue and total variance explained, to determine adequate extracted variables which lead to minimum misclassification rate. The outcomes from these extraction techniques are used to construct the smoothed location models, which then produce two new approaches of classification called 2PCALM and 2DLM. Numerical evidence from simulation studies demonstrates that the computed misclassification rate indicates no significant difference between the extraction techniques in normal and non-normal data. Nevertheless, both proposed approaches are slightly affected for non-normal data and severely affected for highly overlapping groups. Investigations on some real data sets show that the two approaches are competitive with, and better than other existing classification methods. The overall findings reveal that both proposed approaches can be considered as improvement to the location model, and alternatives to other classification methods particularly in handling mixed variables with large binary size.

Keywords: Classification Rule, Smoothed Location Model, Principal Component Analysis, Multiple Correspondence Analysis, Mixed Variables.

Acknowledgement

All praises to the Almighty ALLAH the most Gracious and Merciful, who has blessed me with strength, courage and abundant resources in completing this meaningful journey.

Firstly, I would like to express my utmost gratitude to my supervisor, Dr. Nor Idayu binti Mahat, for her invaluable guidance, support, motivation, comments and fruitful discussions that have led to substantial improvement in this study. I am also thankful for her generosity in sharing knowledge with me and I greatly benefited from her expertise in statistics in particular. Without her valuable support, my thesis would not have been possible. Special appreciation goes to Prof. Qamar Naser Usmani and Prof. Madya Abdull Halim bin Abdul of Universiti Malaysia Perlis (UniMAP), for their willingness in providing hyper computing facilities which have facilitated the simulation process of my study.

My thanks are extended to every staff at the School of Quantitative Sciences, Universiti Utara Malaysia and friends for their direct and indirect helps. In particular, I would like to thank Prof. Dr. Zurni bin Omar (Dean), Dr. Suzilah binti Ismail (Head Department of Mathematics and Statistics) and some senior lecturers such as Prof. Madya Dr. Sharipah Soaad binti Syed Yahaya, Prof. Madya Dr. Haslinda binti Ibrahim and Prof. Madya Dr. Razamin binti Ramli, for providing a conducive learning environment and sharing their profound knowledge on research.

I would like to thank all my friends especially my PhD colleagues, Mrs. Maz Jamilah binti Masnan and Mr. Ch'ng Chee Keong, for interesting discussions and exchange of ideas on relevant topics. Special thanks are also dedicated to Dr. Sharmila binti Karim, Dr. Zakiyah binti Zain, Dr. Adyda binti Ibrahim, Dr. Hasimah binti Sapiri, Dr. Azizah binti Mohd Rohni, Mrs. Norhayati binti Yusof, Dr. Shamsudin bin Ibrahim and Mr. Rusdi @ Indra Zuhdi bin Murat, for their support and friendship throughout my journey.

This study would not have been possible without a financial aid. I would like to thank the Ministry of Higher Education for their financial support. Also to the Universiti Utara Malaysia for giving me the opportunity to pursue my PhD program and gave me ample time in completing this work.

Last but not least, I am very grateful to my beloved husband, Mr. Halim Shah bin Md. Shurani and my mother, Mrs. Bibi Hartom binti Akhbar Shaid, for their endless love, unconditional support and encouragement through all these years. I am also thankful to my entire family members especially Mrs. Bibi Ashah binti Mohammad Saal, Mrs. Bibi Sakinah binti Mohammad Saal, Mr. Mohamad bin Mohammad Saal and Mr. Wali Shah bin Mohammad Saal, for their continued support and love. Special acknowledgement and thanks to my mother-in-law, Mrs. Sabariah binti Yusoff, for her sacrifice, understanding and patience in raising both my adorable sons, Iman Shah bin Halim Shah and Naqib Shah bin Halim Shah.

To all of them, I dedicate this work.

Table of Contents

List of Tables

List of Figures

List of Appendix

List of Publications

Hamid, H. & Mahat, N. I. (2013). Using Principal Component Analysis to Extract Mixed Variables for Smoothed Location Model. *Far East Journal of Mathematical Sciences*, *80*(1), 33-54.

Hamid, H. & Mahat, N. I. (2013). Integration of Variable Extraction and LDA for Mixed Variables Classification Problems. *Modern Applied Science*. (Indexed by Scopus - Accepted).

Hamid, H. & Mahat, N. I. (2012). Strategies of Handling Different Variables Reduction for LDA. *Jurnal Sains dan Matematik, 4*(2), 37-48.

Masnan M. J., Zakaria, A., Shakaff, A. Y., Mahat, N. I., Hamid, H., Subari, N. & Junita M. S. (2012). Principal Component Analysis – A Realization of Classification Success in Multi Sensor Data Fusion. In P. Sanguansat (Eds.), *Principal Component Analysis - Engineering Applications* (pp. 1-24). Croatia, Rijeka: InTech Open Access Publisher.

Hamid, H. & Mahat, N. I (2010). *A New Approach for Classifying Large Number of Mixed Variables*. Paper Presented at the International Conference on Computer & Applied Mathematics, 27-30 October, Paris, France.

Hamid, H. & Mahat, N. I (2010). *Some Investigations on the LDA when Handling Large Number of Variables*. Paper Presented at the International Conference of Applied Mathematics (AMIC) & the 6th East Asia Siam Conference (EASIAM), 22-24 June, Kuala Lumpur, Malaysia.

CHAPTER ONE

CLASSIFICATION

1.1 Introduction

Living things such as animals, plants and humans can be organized into groups based on some similar characteristics. The process of organizing things into groups due to differences between them is called a classification system (Bidelman, 1950). In general, classification can be divided into two types: unsupervised classification and supervised classification (Dudoit & Fridlyand, 2003). In unsupervised classification, group memberships of objects are unknown and its memberships are determined from the data. This classification process involves estimating the number of groups and determining the objects for each group. In contrast, in supervised classification, the group memberships are known at a priory. Thus, this process attempts to create a mathematical form which then can be used to explain the basis of classification and to predict a group of future objects.

The knowledge of classification is helpful to support and to build theories related to differences of objects. For example, in epidemiology studies, classification is used to determine the level of health risks for different populations (Kristensen, 1992). Classification and discrimination are often used interchangeably. However, classification is related to the problem of identifying the sub-populations to which new objects belong to, on the basis of a set of measurements that have been made on it (Krzanowski, 2000, p. 330). Meanwhile, discrimination deals with the separation of groups of data (Esbensen et al., 2002, p. 496). In other areas such as machine

The contents of the thesis is for internal user only

REFERENCES

Abdi, H. & Valentin, D. (2007). Multiple Correspondence Analysis. In N. Salkind (Eds.), *Encyclopedia of Measurement and Statistics* (pp. 3-16). Thousand Oaks, CA: Sage.

Abdi, H. & Williams, L. J. (2010). Principal Component Analysis. *Computational Statistics, 2*, 433-459.

Abel, L., Golmard, J. & Mallet, A. (1993). An Autologistic Model for the Genetic Analysis of Familial Binary Data. *American Journal of Human Genetics, 53*(4), 894-907.

Afifi, A. A. & Elashoff. R. M. (1969). Multivariate Two-sample Tests with Dichotomous and Continuous Variables: The Location Model. *The Annals of Mathematical Statistics, 40*, 290-298.

Agrawal, R. & Srikant, R. (1994). Fast Algorithms for Mining Association Rules. In J. B. Bocca, M. Jarke & C. Zaniolo (Eds.), *Proceedings of the 20th International Conference on Very Large Databases*, 12-15 September 1994, Santiago de Chile, Chile (pp. 487-499). San Francisco, CA: Morgan Kaufmann.

Aitchison, J. & Aitken, C. G. G. (1976). Multivariate Binary Discrimination by the Kernel Method. *Biometrika, 63*, 413-420.

Aktürk, D., Gün, S. & Kumuk, T. (2007). Multiple Correspondence Analysis Technique Used in Analyzing the Categorical Data in Social Sciences. *Journal of Applied Sciences, 7*(4), 585-588.

Alsberg, B. K., Goodacre, R., Rowland, J. J. & Kell, D. B. (1997). Classification of Pyrolysis Mass Spectra by Fuzzy Multivariate Rule Induction-Comparison with Regression, *K*-nearest neighbour, Neural and Decision-tree Methods. *Analytica Chimica Acta, 348*(1-3), 389-407.

Ambroise, C. & McLachlan, G. J. (2002). Selection Bias in Gene Extraction on the Basis of Microarray Gene-Expression Data. *Proceedings of the National Academy of Sciences, 99*(10), 6562-6566.

Amidan, B. G. & Hagedorn, D. N. (1998). *Logistic Regression Applied to Seismic Discrimination* (Report No. PNNL-RR-98-12031). Richland, WA: Pacific Northwest National Laboratory.

An, J. & Chen, Yi-P. P. (2009). Finding Rule Groups to Classify High Dimensional Gene Expression Datasets. *Computational Biology and Chemistry, 33*, 108-113.

An, J., Chen, Yi-P. P. & Chen, H. (2005). DDR: An Index Method for Large Time-Series Datasets. *Information Systems, 30*(5), 333-348.

Andersen, E. B. (1990). *Statistical Analysis of Categorical Data*. New York, NY: Springer-Verlag.

Anderson, J. A. (1972). Separate Sample Logistic Discrimination. *Biometrika*, *59*(1), 19-35.

Anderson, J. A. (1975). Quadratic Logistic Discrimination. *Biometrika, 62*, 149-154.

Anderson, J. A. (1982). Logistic Discrimination. In P. R. Krishnaiah & L. N. Kanal (Eds.), *Handbook of Statistics* (pp. 169-191). Amsterdam, North-Holland.

Anderson, J. A. & Richardson, S. C. (1979). Logistic Discrimination and Bias Correction in Maximum Likelihood Estimation. *Technometrics, 21,* 71-78.

Anderson, T. W. (1958). *An Introduction to Multivariate Statistical Analysis*. New York, NY: John Wiley & Sons, Inc.

Antoniadis, A., Lambert-Lacroix, S. & Leblanc, F. (2003). Effective Dimension Reduction Methods for Tumor Classification using Gene Expression Data. *Bioinformatics, 19*, 563-570.

Arabie, P. & Hubert, L. (1994). Cluster Analysis in Marketing Research. In R. P. Bagozzi (Eds.), *Advanced Methods of Marketing Research* (pp. 160-189). Oxford: Blackwell.

Asan, Z. & Greenacre, M. J. (2011). Biplots of Fuzzy Coded Data. *Fuzzy Sets and Systems, 183*(1), 57-71.

Asparoukhov, O. & Krzanowski, W. J. (2000). Non-parametric Smoothing of the Location Model in Mixed Variable Discrimination. *Statistics and Computing, 10*(4), 289-297.

Asparoukhov, O. & Krzanowski, W. J. (2001). A Comparison of Discriminant Procedures for Binary Variables. *Computational Statistics & Data Analysis, 38*, 139-160.

Auria, L. & Moro, R. A. (2008). *Support Vector Machines (SVM) as a Technique for Solvency Analysis* (Working Paper No. WP-08-811). Berlin: German Institute for Economic Research. Retrieved from http://www.diw.de/documents/publikationen/73/diw_01.c.88369.de/dp811.pdf.

Babbie, E. R. (2009). *The Practice of Social Research* (12th ed.). Belmont, CA: Cengage Learning.

Baccini, P. & Bader, H. P. (1996). *Regionaler Stoffhaushalt: Erfassung, Bewertung und Steuerung* [Regional Material Household: Detecting, Assessment and Controlling]. Heidelberg, Berlin: Spektrum Akademischer Verlag.

Baeka, J. & Kimb, M. (2004). Face Recognition using Partial Least Squares Components. *Pattern Recognition, 37*(6), 1303-1306.

Bar-Hen, A. (2002). Generalized Principal Component Analysis of Continuous and Discrete Variables. *InterStat Statistics, 8*(6), 1-26.

Bar-Hen, A. & Daudin, J. J. (1995). Generalization of the Mahalanobis Distance in the Mixed Case. *Journal of Multivariate Analysis*, *53*(2), 332-342.

Bartholomew, D. J. (1984). Scaling Binary Data using a Factor Model. *Journal of the Royal Statistical Society, Series B (Methodological), 46*(1), 120-123.

Bartholomew, D. J. (1987). *Latent Variable Models and Factor Analysis*. New York, NY: Oxford University Press.

Beh, E. J. (2004). Simple Correspondence Analysis: A Bibliographic Review. *International Statistical Review, 72*(2), 257-284.

Belhumeur, P. N., Hespanha, J. P. & Kriegman, D. J. (1997). Eigenfaces vs. Fisherfaces: Recognition using Class Specific Linear Projection. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 19*(7), 711-720.

Bellman, R. (1961). *Adaptive Control Processes: A Guided Tour*. New Jersey: Princeton University Press.

Benzécri, J. P. (1973). *L'analyse des Données: l'analyse des Correspondances* [Data Analysis: Correspendence Analysis]. Paris: Dunod.

Benzécri, J. P. (1977a). Histoire et Préhistoire de l'analyse des Données: l'analyse des Correspondances [History and Prehistory of Data Analysis: Correspondence Analysis]. *Les Cahiers de l'analyse des Données, 2*, 9-53.

Benzécri, J. P. (1977b). Sur l'analyse des Tableaux Binaires Associés à une Correspondance Multiple [The Analysis of Boolean Tables Associated with a Multiple Correspondence]. *Les Cahiers de l'analyse des Données, 2*, 55-71.

Benzécri, J. P. (1979). Sur le Calcul des Taux D'inertie dans l'analyse d'un Questionnaire. *Les Cahiers de l'Analyse des Données*, *3*, 55-71.

Benzécri, J. P. (1992). *Correspondence Analysis Handbook*. New York, NY: Marcel Dekker, Inc.

Berry, M. W., Dumais, S. T. & O'Brien, G. W. (1995). Using Linear Algebra for Intelligent Information Retrieval. *Society for Industrial and Applied Mathematics* (*SIAM) Review, 37*(4), 573-595.

Bidelman, W. P. (1950). Spectral Classification of Stars Listed in Miss Payne's Catalogue of Stars. *Astrophysical Journal, 113*, 304-316.

Bishop, C. M. (1995). *Neural Networks for Pattern Recognition*. New York, NY: Oxford University Press.

Bittencourt, H. R. & Clarke, R. T. (2003). Logistic Discrimination between Classes with Nearly Equal Spectral Response in High Dimensionality. In *Proceedings of the IEEE International of Geoscience and Remote Sensing Symposium, 6*, 21-25 July 2003, Toulouse, France (pp. 3748-3750). Piscataway, NJ: IEEE Operations Center.

Bittencourt, H. R., Moraes, D. A. O. & Haertel, V. (2007). Single and Multiple Stage Classifiers Implementing Logistic Discrimination. In *Proceedings of the XIII Brazilian of Remote Sensing Symposium, 13*, 21-26 April 2007, Brazil (pp. 6431-6436). Florianopolis, Brazil: National Institute for Space Research (INPE).

Blasius, J. & Thiessen, V. (2000). Methodological Artifacts in Measures of Political Efficacy and Trust: A Multiple Correspondence Analysis. *Political Analysis, 9*(1), 1-20.

Bloomfield, P. (1974). Linear Transformations for Multivariate Binary Data. *Biometrics*, *30*(4), 609-617.

Borgognone, M. G., Bussi, J. & Hough, G. (2001). Principal Component Analysis: Covariance or Correlation Matrix. *Food Quality and Preference, 12*, 323-326.

Boser, B. E., Guyon, I. M. & Vapnik, V. N. (1992). A Training Algorithm for Optimal Margin Classifiers. *Proceedings of the Fifth Annual Workshop on Computational Learning Theory* (pp. 144-152). New York, NY: ACM Press.

Bouguila, N. (2010). On Multivariate Binary Data Clustering and Feature Weighting. *Computational Statistics and Data Analysis, 54*, 120-134.

Boulesteix, A. (2004). PLS Dimension Reduction for Classification with Microarray Data. *Statistical Applications in Genetics and Molecular Biology, 3*, 1-33.

Braga-Neto, U. M. & Dougherty, E. R. (2004). Is Cross-validation Valid for Small-sample Microarray Classification? *Bioinformatics, 20*(3), 374-380.

Braga-Neto, U. M., Hashimoto, R., Dougherty, E. R., Nguyen, D. V. & Carroll, R. J. (2004). Is Cross-validation Better than Resubstitution for Ranking Genes? *Bioinformatics*, *20*(2), 253-258.

Breiman, L. (1996). Bagging Predictors. *Machine Learning, 26*, 123-140.

Breiman, L. (2001). Random Forests. *Machine Learning, 45*, 5-32.

Breiman, L., Friedman, J. H., Olshen, R. A. & Stone, C. J. (1984). *Classification and Regression Trees*. Belmont, CA: Wadsworth.

Browne, R. P. & McNicholas, P. D. (2012). Model-based Clustering, Classification, and Discriminant Analysis of Data with Mixed Type. *Journal of Statistical Planning and Inference, 142*, 2976-2984.

Buja, A. (1990). Remarks on Functional Canonical Variates, Alternating Least Squares Methods and Ace. *The Annals of Statistics, 18*(3), 1032-1069.

Bull, S. B. & Donner, A. (1987). The Efficiency of Multinominal Logistic Regression compared with Multiple Group Discriminant Analysis. *Journal of the American Statistical Association, 82*, 1118-1122.

Bura, E. & Pfeiffer, R. M. (2003). Graphical Methods for Class Prediction using Dimension Reduction Techniques on DNA Microarray Data. *Bioinformatics, 19*(10), 1252-1258.

Burges, C. J. C. (1998). A Tutorial on Support Vector Machines for Pattern Recognition. *Data Mining and Knowledge Discovery, 2*(2), 121-167.

Burt, C. (1950). The Factorial Analysis of Qualitative Data. *British Journal of Psychology, 3*, 166-185.

Burt, C. (1953). Scale Analysis and Factor Analysis. *British Journal of Statistical Psychology, 6*, 5-23.

Buttrey, S. E. (1998). Nearest-Neighbor Classification with Categorical Variables. *Computational Statistics & Data Analysis, 28*(2), 157-169.

Cadima, J. & Jolliffe, I .T. (2001). Variable Selection and the Interpretation of Principal Subspaces. *Journal of Agricultural, Biological and Environmental Statistics, 6*(1), 62-79.

Cadima, J., Cerdeira, J. O. & Minhoto, M. (2004). Computational Aspects of Algorithms for Variable Selection in the Context of Principal Components. *Computational Statistics & Data Analysis, 47*, 225-236.

Cai, D., He, X., Zhou, K., Han, J. & Bao, H. (2007). Locality Sensitive Discriminant Analysis. *Proceedings of the 20th International Joint Conference on Artificial Intelligence* (pp. 708-713). San Francisco, CA: Morgan Kaufmann.

Camiz, S. & Gomes, G. C. (2013). Joint Correspondence Analysis versus Multiple Correspondence Analysis: A Solution to an Undetected Problem. In A. Giusti,

G. Ritter & M. Vichi (Eds.), *Classification and Data Mining: Studies in Classification, Data Analysis, and Knowledge Organization* (pp. 11-19). Berlin, Heidelberg: Springer-Verlag.

Caprihan, A., Pearlson, G. D. & Calhoun, V. D. (2008). Application of Principal Component Analysis to Distinguish Patients with Schizophrenia from Healthy Controls based on Fractional Anisotropy Measurements. *Neuroimage, 42*(2), 675-682.

Carroll, J. D. & Green P. E. (1988). An INDSCAL-based Approach to Multiple Correspondence Analysis. *Journal of Marketing Research, 25*, 193-203.

Cattell, R. B. (1966). The Scree Test for the Number of Factors. *Multivariate Behavioral Research, 1*, 245-276.

Cerioli, A., Riani, M. & Atkinson, A. C. (2006). Robust Classification with Categorical Variables. In A. Rizzi & M. Vichi (Eds.), *Proceedings in Computational Statistics* (pp. 507 -519). Berlin, Heidelberg: Springer-Verlag.

Chakravarti, I. M., Laha, R. G. & Roy, J. (1967). *Handbook of Methods of Applied Statistics* (Vol. 1). New York, NY: John Wiley & Sons, Inc.

Chan, L-H., Salleh, Sh-H., Ting, C-M. & Ariff, A. K. (2008, August). Face Identification and Verification using PCA and LDA. *International Symposium on Information Technology, 2*, 1-6.

Chan, Y. B. & Hall, P. (2009). Scale Adjustments for Classifiers in High-Dimensional, Low Sample Size Settings. *Biometrika, 96*(2), 469-478.

Chan, Y. H. (2005). Biostatistics 303: Discriminant Analysis. *Singapore Medical Journal, 46*(2), 54-61.

Chandan, M., White, H. & Wuyts, M. (1998). *Econometrics and Data Analysis for Developing Countries*. London: Routledge.

Chang, P. C. & Afifi, A. A. (1974). Classification based on Dichotomous and Continuous Variables. *Journal of the American Statistical Association, 69*(346), 336- 39.

Charu, C. A. (2001). On The Effects of Dimensionality Reduction on High Dimensional Similarity Search. *Proceedings of the Twentieth ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems* (pp. 256-266). New York, NY: ACM Press.

Chen, L. F., Liao, H. Y. M., Ko, M. T., Lin, J. C. & Yu, G. J. (2000). A New LDA-based Face Recognition System which Can Solve the Small Sample Size Problem. *Pattern Recognition, 33*(10), 1713-1726.

Chen, Z., Tian, L. & Geng, Z. (2008). Support Vector Machine Classifier with Feature Extraction by Kernel PCA for Intrusion Detection. *Journal of Information and Computational Science, 5*(6), 2495-2508.

Cheng, W. (2005). *Factor Analysis for Stock Performance* (Master's thesis, Worcester Polytechinic Institute). Retrieved from http://www.wpi.edu/Pubs/ETD/Available/etd-050405-180040/unrestricted/Wei_Cheng.pdf.

Chiaromonte, F. & Martinelli, J. (2002). Dimension Reduction Strategies for Analyzing Global Gene Expression Data with a Response. *Mathematical Biosciences, 176*(1), 123-144.

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.

Chou, Y-T. & Wang, W-C. (2010). Checking Dimensionality in Item Response Models with Principal Component Analysis on Standardized Residuals. *Educational and Psychological Measurement, 70*(5), 717-731.

Christoffersson, A. (1975). Factor Analysis of Dichotomized Variables. *Psychometrika, 40*(1), 5-32.

Cochran, W. G. & Hopkins, C. (1961). Some Classification Methods with Multivariate Qualitative Data. *Biometrics, 17*, 10-32.

Collins, M., Dasgupta, S. & Schapire, R. E. (2001). A Generalization of Principal Component Analysis to the Exponential Family. In T. G. Dietterich, S. Becker & Z. Ghahramani (Eds.), *Proceedings of the Advances in Neural Information Processing Systems, 14*, 3-8 December 2001, Canada (pp. 617-624). MIT Press.

Cook, R. D., Buja, A. & Cabrera, J. (1993). Projection Pursuit Indexes Based on Orthonormal Function Expansions. *Journal of Computational and Graphical Statistics, 2,* 225-250.

Cook, R. D. (1998). *Regression Graphics*. New York, NY: John Wiley & Sons, Inc.

Cook, R. D. & Lee, H. (1999). Dimension Reduction in Binary Response Regression. *Journal of the American Statistical Association, 94*(448), 1187-1200.

Cooley, W. W. & Lohnes, P. R. (1971). *Multivariate Data Analysis*. New York, NY: John Wiley & Sons, Inc.

Copeland, K. T., Checkoway, H., McMichael, A. J. & Holbrook, A. (1977). Bias due to Misclassification in the Estimation of Relative Risk. *American Journal of Epidemiology, 105*(5), 488-495.

Costello, A. B. & Osborne, J. W. (2005). Best Practices in Exploratory Factor Analysis. *Practical Assessment Research & Evaluation, 10*(7), 272-280.

Cox, D. R. (1966). Some Procedures Associated with the Logistic Qualitative Response Curve. In J. Neyman & F. N. David (Eds.), *Research Papers in Statistics* (pp. 55-71). New York, NY: John Wiley & Sons, Inc.

Cox, D. R. (1972). The Analysis of Multivariate Binary Data. *Applied Statistics, 21*(2), 113-120.

Cox, T. F. & Pearce, K. F. (1997). A Robust Logistic Diascrimination Model. *Statistics and Computing, 7*(3), 155-161.

Cristianini, N. & Taylor, J. S. (2000). *An Introduction to Support Vector Machines and Other Kernel-based Learning Methods*. Cambridge, MA: University Press.

Cuadras, C. M. (1992). Some Examples of Distance based Discrimination. *Biometrical Letters, 29*, 3-20.

Cuadras, C. M., Fortiana, J. & Oliva, F. (1997). The Proximity of an Individual to a Population with Applications in Discriminant Analysis. *Journal of Classification, 14*, 117-136.

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.

Daniels, M. J. & Kass, R. E. (2001). Shrinkage Estimators for Covariance Matrices. *Biometrics, 57*, 1173-1184.

Das, K. (2007). *Feature Extraction and Classification for Large-scale Data Analysis* (Unpublished doctoral dissertation). Department of Electrical and Computer Engineering, University of California, Irvine, USA.

Das, K., Meyer, J. & Nenadic, Z. (2006). Analysis of Large-Scale Brain Data for Brain- Computer Interfaces. In *Proceedings of the 28th IEEE Annual International Conference of Engineering in Medicine and Biology Society*, 30 August - 3 September 2006, New York (pp. 5731-5734). Retrieved from http://cbmspc.eng.uci.edu/PUBLICATIONS/zn:06b.pdf.

Das, K., Osechinskiy, S. & Nenadic, Z. (2007). A Classwise PCA-based Recognition of Neural Data for Brain-Computer Interfaces. In *Proceedings of the 29th IEEE Annual International Conference of Engineering in Medicine and Biology Society*, 22-26 August 2007, France (pp. 6519-6522). Retrieved from http://cbmspc.eng.uci.edu/PUBLICATIONS/zn:07c.pdf.

Daudin, J. J. (1986). Selection of Variables in Mixed-variable Discriminant Analysis. *Biometrics, 42*(3), 473-481.

Daudin, J. J. & Bar-Hen, A. (1999). Selection in Discriminant Analysis with Continuous and Discrete Variables. *Computational Statistics and Data Analysis, 32*(2), 161-175.

Daudin, J. J. & Trecourt, P. (1980). Analyse Factorielle des Correspondences et Modèle Log-linéaire: Comparaison des deux Méthodes sur un Exemple [Correspondence Analysis and log-linear Model: Comparison of Both Models on an Example]. *Revue de Statistique Appliquée, 28*, 5-24.

Debruyne, M. (2009). An Outlier Map for Support Vector Machine Classification. *Annals of Applied Statistics, 3*(4), 1566-1580.

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K. & Harshman, R. (1990). Indexing by Latent Semantic Analysis. *Journal of the American Society for Information Science, 41*(6), 391-407.

de Leeuw, J. (1984). *Statistical Properties of Multiple Correspondence Analysis*. Paper Presented at the New Multivariate Methods in Statistics Conference: The 1984 Joint Summer Research Conference Series in Mathematical Sciences, 10-16 June, Brunswick, Maine.

de Leeuw, J. (1987). Nonlinear Multivariate Analysis with Optimal Scaling. In P. Legendre & L. Legendre (Eds.), *Developments in Numerical Ecology*. Berlin, Heidelberg: Springer-Verlag.

de Leeuw, J. (1998). Here's Looking at Multivaribles. In J. Blasius & M. J. Greenacre (Eds.), *Visualization of Categorical Data* (pp. 1–11). San Diego: Academic Press.

de Leeuw, J. & Mair, P. (2009). Simple and Canonical Correspondence Analysis Using the R Package anacor. *Journal of Statistical Software, 31*(5), 1-18.

Deng, H. B., Jin, L. W., Zhen, L. X. & Huang, J. C. (2005). A New Facial Expression Recognition Method on Local Gabor Filter Bank and PCA plus LDA. *International Journal of Information Technology, 11*(11), 86-96.

D'Enza, A. I. & Greenacre, M. J. (2012). Multiple Correspondence Analysis for the Quantification and Visualization of Large Categorical Data Sets. In A. Di Ciaccio, M. Coli & J. M. A. Ibaňez (Eds.), *Advanced Statistical Methods for the Analysis of Large Data-Sets: Studies in Theoretical and Applied Statistics* (pp. 453-463). Berlin, Heidelberg: Springer-Verlag.

Devivjer, P. & Kittler, J. (1982). *Pattern Recognition: A Statistical Approach*. Englewood Cliffs, NJ: Prentice Hall.

Dillon, W. R. & Goldstein, M. (1978). On the Performance of Some Multinomial Classification Rules. *Journal of the American Statistical Association, 73*, 305-313.

Dillon, W. R. & Goldstein, M. (1984). *Multivariate Analysis: Methods and Applications*. New York, NY: John Wiley & Sons, Inc.

DiPillo, P. J. (1976). The Application of Bias to Discriminant Analysis. *Communications in Statistics - Theory and Methods, 5*(9), 843-854.

Doey, L. & Kurta, J. (2011). Correspondence Analysis Applied to Psychological Research. *Quantitative Methods for Psychology, 7*(1), 5-14.

Dom, B. E. (1997). MDL Estimation for Small Samples Sizes and Its Application to Segmentation Binary Strings. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. 280 - 287.

Domeniconi, C., Peng, J. & Gunopulos, D. (2002). Locally Adaptive Metric Nearest-neighbor Classification. *IEEE Transactions on Pattern Analysis and Machine Intelligent, 24*(9), 1281-1285.

Doumpos, M. & Zopounidis, C. (2002). *Multicriteria Decision Aid Classification Methods*. Dordrecht, Netherlands: Kluwer Academic Publishers.

Duan, N. & Li, K. C. (1991). Slicing Regression: A Link-free Regression Method. *The Annals of Statistics, 19*, 505-530.

Dubois, J. P. & Abdul-Latif, M. (2005). Improved *M*-ary Signal Detection using Support Vector Machine Classifiers. *Proceedings of World Academy of Science, Engineering and Technology, 7*, 264-268.

Duda, R. O., Hart, P. E. & Stork, D. G. (2001). *Pattern Classification* (2nd ed.). New York, NY: John Wiley & Sons, Inc.

Dudoit, S. & Fridlyand, J. (2003). Introduction to Classification in Microarray Experiments. In D. P. Berrar, W. Dubitzky & M. Granzow (Eds.), *A Practical Approach to Microarray Data Analysis* (pp. 132-149). Massachusetts, USA: Kluwer Academic Publishers.

Dudoit, S., Fridlyand, J. & Speed, T. P. (2002). Comparison of Discriminant Methods for the Classification of Tumors using Gene Expressing Data: Applications and Case Studies. *Journal of the American Statistical Association, 97*(457), 77-87.

Eastment, H. T. & Krzanowski, W. J. (1982). Cross-Validation Choice of the Number of Components from a Principal Component Analysis. *Technometrics, 24*, 73-77.

Edward, J. (1991). *A User's Guide to Principal Components*. New York, NY: John Wiley & Sons, Inc.

Efron, B. (1975). The Efficiency of Logistic Regression Compared to Normal Discriminant Analysis. *Journal of the American Statistical Association, 70,* 892-898.

Efron, B. (1979). Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics, 7,* 1-26.

Elisseeff, A. & Pontil, M. (2003). Leave-one-out Error and Stability of Learning Algorithms with Applications. In J. Suykens, G. Horvath, S. Basu, C. Micchelli, & J. Vandewalle (Eds.), *Advances in Learning Theory: Methods, Models and Applications* (pp. 111-124). NATO Science Series III: Computer & Systems Sciences (Vol. 190). Amsterdam: IOS Press.

Esbensen, K. H., Guyot, D., Westad, F. & Houmoller, L. P. (2002). *Multivariate Data Analysis in Practice: An Introduction to Multivariate Data Analysis and Experimental Design* (5th ed.). Oslo, Norway: CAMO AS.

Etemad, K. & Chellapa, R. (1997). Discriminant Analysis for Recognition of Human Face Images. *Journal of Optical Society of American, 8,* 1724-1733.

Everitt, B. S. (1988). A Finite Mixture Model for the Clustering of Mixed-mode Data. *Statistics and Probability Letters, 6,* 305-309.

Everitt, B. S. & Merette, C. (1990). The Clustering of Mixed-mode Data: A Comparison of Possible Approaches. *Journal of Applied Statistics, 17,* 283-297.

Ewans, W. J. & Grant, G. R. (2001). *Statistical Methods in Bioinformatics: An Introduction*. New York, NY: Springer-Verlag.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C. & Strahan, E. J. (1999). Evaluating the Use of Exploratory Factor Analysis in Psychological Research. *Psychological Methods, 4*(3), 272-299.

Farmer, S. A. (1971). An Investigation into the Results of Principal Component Analysis of Data Derived from Random Numbers. *Statistician, 20,* 63-72.

Fisher R. A. (1936). The Use of Multiple Measurements in Taxonomic Problems. *Annals of Eugenics, 7*(2), 179-188.

Fix, E. & Hodges, J. L. (1951). *Discriminatory Analysis - Nonparametric Discrimination: Consistency Properties* (Report No. 51-4). Retrieved from http://www.dtic.mil/dtic/tr/fulltext/u2/a800276.pdf.

Fodor, I. K. (2002). *A Survey of Dimension Reduction Techniques* (Report No. LLNL-RR-02-148494). California, CA: Lawrence Livermore National Laboratory.

Ford, I., Norris, J. & Ahmadi, S. (1995). Model Inconsistency, Illustrated by the Cox Proportional Hazards Model. *Statistics in Medicine, 14*, 735-746.

Franklin, S. B., Gibson, D. J., Robertson, P. A., Pohlmann, J. T. & Fralish, J. S. (1995). Parallel Analysis: A Method for Determining Significant Principal Components. *Journal of Vegetation Science, 6*(1), 99-106.

Fränti, P., Xu, M. & Kärkkäinen, I. (2003). Classification of Binary Vectors by Using *Δ*SC Distance to Minimize Stochastic Complexity. *Pattern Recognition Letters, 24*(1), 65-73.

Friedman, J. H. (1989). Regularized Discriminant Analysis. *Journal of the American Statistical Association, 84*(405), 165-175.

Fukunaga, K. (1990). *Introduction to Statistical Pattern Recognition* (2^{nd} ed.). San Diego, CA: Academic Press.

Gail, M. H., Wienand, S. & Piantadosi, S. (1984). Biased Estimates of Treatment Effects in Randomized Experiments with Nonlinear Regressions and Omitted Variables. *Biometrika, 71*, 431-444.

Garcia, T. & Grande, I. (2003). A Model for the Valuation of Farmland in Spain: The Case for the Use of Multivariate Analysis. *Journal of Property Investment & Finance, 21*(2), 136-153.

Garthwaite, P. M. (1994). An Interpretation of Partial Least Squares. *Journal of the American Statistical Association, 89*(425), 122-127.

Gascuel, O. & Caraux, G. (1992). Distribution-free Performance Bounds with the Resubstitution Error Estimate. *Pattern Recognition Letters, 13*, 757-764.

Ghosh, A. (2011). Forecasting BSE Sensex under Optimal Conditions: An Investigation Post Factor Analysis. *Journal of Business Studies Quarterly, 3*(2), 57-73.

Ghosh, D. (2002). Singular Value Decomposition Regression Models for Classification of Tumors from Microarray Experiments. In R. B. Altman, A. K. Dunker, L. Hunter, K. Lauderdale & T. E. Klein (Eds.), *Proceedings of the 2002 Pacific Symposium on Biocomputing* (pp. 18-29). Kauai, Hawaii: World Scientific.

Gibson, A. R., Baker, A. J. & Moeed, A. (1984). Morphometric Variation in Introduced Populations of the Common Myna (*Acridotheres Tristis*): An Application of the Jackknife to Principal Component Analysis. *Systematic Biology*, *33*(4), 408-421.

Giordano, F. R., Fox, W. P., Horton, S. B. & Weir, M. D. (2009). *A First Course in Mathematical Modeling* (4th ed.). CA: Belmont. Cengage Learning.

Gifi, A. (1990). *Nonlinear Multivariate Analysis*. Chichester: John Wiley & Sons, Inc.

Giri, N. C. (2004). *Multivariate Statistical Analysis*. New York, NY: Marcel Dekker, Inc.

Girolami, M. (2001). The Topographic Organization and Visualization of Binary Data using Multivariate-Bernoulli Latent Variable Models. *IEEE Transactions on Neural Networks, 12*(6), 1367-1374.

Glick, N. (1978). Additive Estimators for Probabilities of Correct Classification. *Pattern Recognition, 10*, 211-222.

Glynn, D. (2012). Correspondence Analysis: Exploring Data and Identifying Patterns. In D. Glynn & J. Robinson (Eds.), *Polysemy and Synonymy: Corpus Methods and Applications in Cognitive Linguistics* (pp. 133-179). Amsterdam: John Benjamins.

Gnanadesikan, R. (1977). *Methods for Statistical Data Analysis of Multivariate Observations*. New York, NY: John Wiley & Sons, Inc.

Gnanadesikan, R., Roger, K., Breiman, L., Dunn, O. J., Friedman, J. H., Fu, K. S., Hartigan, J. A., Kettenring, J. R., Lachenbruch, P. A., Olshen, R. A. & Rohlf, F. J. (1989). Discriminant Analysis and Clustering: Panel on Discriminant Analysis, Classification and Clustering. *Statistical Science*, *4*(1), 34-69.

Goodstein, R. E. (1987). An Introduction to Discriminant Analysis. *Journal of Research in Music Education, 35*(1), 7-11.

Gower, J. C. (1966). Some Distance Properties of Latent Root and Vector Methods Used in Multivariate Analysis. *Biometrika, 53*(3-4), 325-338.

Gower, J. C. (1971). A General Coefficient of Similarity and Some of Its Properties. *Biometrics, 27*, 857-874.

Green, P. E., Krieger, A. M. & Carroll, J. D. (1987). Multidimensional Scaling: A Complementary Approach. *Journal of Advertising Research*, 21-27.

Greenacre, M. J. (1984). *Theory and Applications of Correspondence Analysis.* London: Academic Press.

Greenacre, M. J. (1998). Diagnostics for Joint Displays in Correspondence Analysis. In J. Blasius & M. J. Greenacre (Eds.), *Visualization of Categorical Data* (pp. 221-238). New York, NY: Academic Press.

Greenacre, M. J. (2005). From Simple to Multiple and Joint Correspondence Analysis. In M. J. Greenacre & J. Blasius (Eds.), *Multiple Correspondence Analysis and Related Methods* (pp. 41-76). London: Chapman and Hall/CRC.

Greenacre, M. J. (2006). Tying Up the Loose Ends in Simple, Multiple and Joint Correspondence Analysis. In A. Rizzi & M. Vichi (Eds.), *Proceedings in Computational Statistics* (pp. 163-185). Berlin, Heidelberg: Physica-Verlag.

Greenacre, M. J. (2007). *Correspondence Analysis in Practice* (2nd ed.). Boca Raton: Chapman & Hall/CRC.

Greenacre, M. J. (2010). Correspondence Analysis. *Computational Statistics, 2*(5), 613-619.

Greenacre, M. J. & Blasius, J. (Eds.) (2006). *Multiple Correspondence Analysis and Related Methods.* London: Chapman & Hall/CRC.

Greenland, S. (1988). Variance Estimation for Epidemiologic Effect Estimates under Misclassification. *Statistics in Medicine, 7*(7), 745-757.

Greenshtein, E. & Ritov, Y. (2004). Persistency in High Dimensional Linear Predictor-Selection and the Virtue of Over-parametrization. *Bernoulli, 10*, 971-988.

Grossman, G. D., Nickerson, D. M. & Freeman, M. C. (1991). Principal Component Analyses of Assemblage Structure Data: Utility of Tests based on Eigenvalues. *Ecology, 72*(1), 341-347.

Guérif, S. (2008). Unsupervised Variable Selection: When Random Rankings Sound as Irrelevancy. *Journal of Machine Learning Research - New Challenges for Feature Selection in Data Mining and Knowledge Discovery, 4*, 163-177.

Guerreiro, P. M. C. (2008). *Linear Discriminant Analysis Algorithms* (Unpublished master's thesis). Technical University of Lisbon, Portugal.

Guo, Y., Hastie, T. & Tibshirani, R. (2007). Regularized Linear Discriminant Analysis and Its Application in Microarrays. *Biostatistics, 8*(1), 86-100.

Guttman, L. (1941). The Quantification of a Class of Attributes: A Theory and Method of Scale Construction. In P. Horst, P. Wallin & L. Guttman (Eds.), *The Prediction of Personal Adjustment* (pp. 319-348). New York, NY: Social Science Research Council.

Guyon, I. & Elisseeff, A. (2003). An Introduction to Variable and Feature Selection. *Journal of Machine Learning Research, 3*, 1157-1182.

Gyllenberg, M., Koski, T. & Verlaan, M. (1997). Classification of Binary Vectors by Stochastic Complexity. *Journal of Multivariate Analysis, 63*, 47-72.

Habbema, J. D. F., Hermans, J. & van den Broek, K. (1974). A Stepwise Discriminant Analysis Program using Density Estimation. In G. Bruckman (Eds.), *Proceedings in Computational Statistics*, Vienna (pp. 101-110). Heidelberg: Physica-Verlag.

Haertel, V. & Landgrebe, D. (1999). On the Classification of Classes with Nearly Equal Spectral Response in Remote Sensing Hyperspectral Image Data. *IEEE Transactions on Geoscience and Remote Sensing, 37*(5), 2374-2386.

Hair, J. F., Anderson, R. E., Tatham, R. L. & Black, W. C. (1998). *Multivariate Data Analysis* (5th ed.). New Jersey: Prentice-Hall, Inc.

Hall, P. (1981). Optimal Near Neighbour Estimator for Use in Discriminant Analysis. *Biometrika, 68*(2), 572-575.

Han, H., Cao, Z., Gu, B. & Ren, N. (2010). PCA-SVM-based Automated Fault Detection and Diagnosis (AFDD) for Vapor-compression Refrigeration Systems. *Journal of HVAC & R Research, 16*(3), 295-313.

Hand, D. J. (1997). *Construction and Assessment of Classification Rules*. Wiley Series in Probability and Statistics. Chichester: John Wiley & Sons, Inc.

Härdle, W. & Hlávka, Z. (2007). *Multivariate Statistics: Exercises and Solutions*. New York, NY: Springer-Verlag.

Hastie, T. & Tibshirani, R. (1986). Generalized Additive Models. *Statistical Science, 1*(3), 297-318.

Hastie, T., Tibshirani, R. & Friedman, J. H. (2009). *The Elements of Statistical Learning: Data Mining, Inference and Prediction* (2nd ed.). Springer-Verlag.

Hermans, R. & Kulvik, M. (2004). Measuring Intellectual Capital and Sources of Equity Financing - Value Platform Perspective within the Finnish Biopharmaceutical Industry. *International Journal of Learning and Intellectual Capital, 1*(3), 282-303.

Hertz, J. A., Krogh, A. S. & Palmer, R. G. (1991). *Introduction of Theory of Neural Computation*. Addison-Wesley, CA: Elsevier Science.

Hills, M. (1967). Discrimination and Allocation with Discrete Data. *Applied Statistics, 16*, 237-250.

Hirsch, O., Bösner, S., Hüllermeier, E., Senge, R., Dembczynski, K. & Donner-Banzhoff, N. (2011). Multivariate Modeling to Identify Patterns in Clinical Data: The Example of Chest Pain. BMC Medical Research Methodology, *11*, 155-164.

Hirst, D. (1996). Error-rate Estimation in Multiple-Group Linear Discriminant Analysis. *Technometrics, 38*, 389-399.

Hoadley, B. (2001). Comment on "Statistical Modeling: The Two Cultures" by Breiman, L. *Statistical Science, 16*(3), 220-224.

Hoffbeck, J. P. & Landgrebe, D. A. (1996). Covariance Matrix Estimation and Classification with Limited Training Data. *IEEE Transactions on Pattern Analysis and Machine Intelligent, 18*(7), 763-767.

Hoffman, D. L. & Batra, R. (1991). Viewer Response to Programs: Dimensionality and Concurrent Behavior. *Journal of Advertising Research*, 46-56.

Hoffman, D. L. & Franke, G. R. (1986). Corresponding Analysis: Graphical Representation of Categorical Data in Marketing Research. *Journal of Marketing Research, 23*(3), 213-227.

Holland, J. H. (1975). *Adaptation in Natural and Artificial Systems:* An Introductory Analysis with Applications to Biology, Control and Artificial Intelligence. Ann Arbor: The University of Michigan Press.

Hotelling, H. (1933). Analysis of a Complex of Statistical Variables into Principal Components. *Journal of Educational Psychology*, *24*, 417-441.

Hoyle, D. C. (2008). Automatic PCA Dimension Selection for High Dimensional Data and Small Sample Sizes. *Journal of Machine Learning Research, 9,* 2733-2759.

Huang, F. J. & LeCun, Y. (2006). Large-scale Learning with SVM and Convolutional Nets for Generic Object Categorization. *Proceedings of Computer Vision and Pattern Recognition Conference*, 17-22 June 2006 (pp. 284-291). Piscataway, NJ: IEEE Press.

Huang, H. L. & Antonelli, P. (2001). Application of Principal Component Analysis to High-Resolution Infrared Measurement Compression and Retrieval. *Journal of Applied Meteorology, 40*, 365-388.

Huang. R., Liu, Q., Lu, H. & Ma, S. (2002). Solving the Small Sample Size Problem of LDA. In *Proceedings of the 16th International Conference on Pattern Recognition, 3* (pp. 29-32). Washington, DC: IEEE Computer Society Press.

Huang, X. & Pan, W. (2003). Linear Regression and Two-class Classification with Gene Expression Data. *Bioinformatics, 19*(16), 2072-2078.

Huang, W., Nakamori, Y. & Wang, Sh-Y. (2005). Forecasting Stock Market Movement Direction with Support Vector Machines. *Computers & Operations Research, 32*(10), 2513-2522.

Huang, Z., Chen, H., Hsu, Ch.-J., Chen, W-H. & Wu, S. (2004). Credit Rating Analysis with Support Vector Machines and Neural Networks: A Market Comparative Study. *Decision Support Systems, 37*(4), 543-558.

Huber, P. J. (1985). Projection Pursuit. *The Annals of Statistics, 13*(2), 435-475.

Hull, D. (1994). Improving Text Retrieval for the Routing Problem using Latent Semantic Indexing. In *Proceedings of the 17th Annual International ACM-SIGIR Conference on Research and Development in Information Retrieval*, 3-6 July 1994, Dublin, Ireland (pp. 282-291). New York, NY: Springer-Verlag.

Hwang, H., Tomiuk, M. A. & Takane, Y. (2009). Correspondence Analysis, Multiple Correspondence Analysis and Recent Developments. In R. E. Millsap & A. Maydeu-Olivares (Eds.), *The SAGE Handbook of Quantitative Methods in Psychology* (pp. 243-263). Thousand Oaks: Sage.

Hwang, W., Kim, T-k. & Kee, S-C. (2004). *LDA with Subgroup PCA Method for Facial Image Retrieval*. Paper Presented at the 5th International Workshop on Image Analysis for Multimedia Interactive Services, 21-23 April, Lisbon, Portugal.

Ibragimov, I. A. & Zaitsev, A. I. (1996). *Probability Theory and Mathematical Statistics*. Amsterdam: Gordon & Breach Publishers.

Illowsky, B. & Dean, S. (2010). *Collaborative Statistics*. Texas: Maxfield Foundation.

Jackson, D. A. (1993). Stopping Rules in Principal Components Analysis: A Comparison of Heuristical and Statistical Approaches. *Ecology, 74*(8), 2204-2214.

Jackson, J. E. (1991). *A User's Guide to Principal Components*. New York, NY: John Wiley & Sons, Inc.

Jaume, P-S., Darío, M-I. & Fernando, D-de-M. (2006). *Support Vector Machines for Continuous Speech Recognition*. Paper Presented at the 14th European Signal Processing Conference, 4-8 September, Florence, Italy.

Jeffers, J. N. R. (1967). Two Case Studies in the Application of Principal Component Analysis. *Applied Statistics, 16*(3), 225-236.

Jenkins, L. & Anderson, M. (2003). A Multivariate Statistical Approach to Reducing the Number of Variables in Data Envelopment Analysis. *European Journal of Operational Research, 147*(1), 51-61.

Jiang, W. & Simon, R. (2007). A Comparison of Bootstrap Methods and an Adjusted Bootstrap Approach for Estimating Prediction Error in Microarray Classification. *Statistics in Medicine, 26*, 5320-5334.

Joachims, T. (1998). Making Large-Scale Support Vector Machine Learning Practical. In B. Schölkopf, C. J. C. Burges & A. J. Smola (Eds.), *Advances in Kernel Methods - Support Vector Learning* (pp. 169-184). Cambridge, MA: MIT Press.

Joe, H. (2006). Generating Random Correlation Matrices based on Partial Correlations. *Journal of Multivariate Analysis, 97*, 2177-2189.

Johnson, R. A. & Wichern, D. W. (1992). *Applied Multivariate Statistical Analysis* (3rd ed.). New Jersey: Prentice Hall, Inc.

Jolliffe, I. T. (1972). *Discarding Variables in a Principal Component Analysis I: Artificial Data*. *Applied Statistics, 21*(2), 160-173.

Jolliffe, I. T. (1986). *Principal Component Analysis*. New York, NY: Springer-Verlag.

Jolliffe, I. T. (2002). *Principal Component Analysis* (2nd ed.). New York, NY: Springer-Verlag.

Juan, A. & Vidal, E. (2002). On The Use of Bernoulli Mixture Models for Text Classification. *Pattern Recognition, 35*, 2705-2710.

Kabán, A. & Girolami, M. A. (2002). Fast Extraction of Semantic Features from a Latent Semantic Indexed Corpus. *Neural Processing Letters, 15*(1), 31-43.

Kaciak, E. & Louviere, J. (1990). Multiple Correspondence Analysis of Multiple Choice Data. *Journal of Marketing Research, 27*, 455-465.

Kaiser, H. F. (1960). The Application of Electronic Computers to Factor Analysis. *Educational and Psychological Measurement, 20*, 141-151.

Kaminska, A., Ickowicz, A., Plouin, P., Bru, M. F., Dellatolas, G. & Dulac, O. (1999). Delineation of Cryptogenic Lennox–Gastaut Syndrome and Myoclonic Astatic Epilepsy using Multiple Correspondence Analysis. *Epilepsy Research, 36*, 15-29.

Karacaoren, B. & Kadarmideen, H. N. (2008). Principal Component and Clustering Analysis of Functional Traits in Swiss Dairy Cattle. *Turkey Journal of Veterinary and Animal Sciences, 32*, 163-171.

Katz, M. H. (2006). *Multivariate Analysis: A Practical Guide for Clinicians* (2nd ed.). Cambridge: Cambridge University Press.

Kearns, M. (1997). A Bound on the Error of Cross Validation using the Approximation and Estimation Rates, With Consequences for the Training-Test Split. *Neural Computation, 9*, 1143-1161.

Kerschen, G. & Golinval, J. C. (2002). Non-Linear Generalization of Principal Component Analysis: From A Global to A Local Approach. *Journal of Sound and Vibration, 254*(5), 867-876.

Kim, H. (1992). Measures of Influence in Correspondence Analysis. *Journal of Statistical Computation and Simulation, 40*(3-4), 201-217.

Kim, H. C., Kim, D. & Bang, S. Y. (2003). Extensions of LDA by PCA Mixture Model and Class-wise Features. *Journal of the Pattern Recognition Society, 36*, 1095-1105.

Kim, J. O. & Mueller, C. W. (1978). *Factor Analysis: Statistical Methods and Practical Issues*. Beverly Hills, CA: Sage.

Kim, K-j. (2003). Financial Time Series Forecasting using Support Vector Machines. *Neurocomputing, 55*(1-2), 307-319.

Kim, M. S., Whang, K-Y. & Moon, Y-S. (2012). Horizontal Reduction: Instance-Level Dimensionality Reduction for Similarity Search in Large Document Databases. In *Proceedings of the 28th International Conference on Data Engineering*, 1-5 April 2012 (pp. 1061-1072). IEEE Computer Society Press.

King, J. R. & Jackson, D. A. (1999). Variable Selection in Large Environmental Data Sets using Principal Components Analysis. *Environmetrics, 10*, 67-77.

Klawun, C. & Wilkins, C. L. (1996). Joint Neural Network Interpretation of Infrared and Mass Spectra. *Journal of Chemical Information and Computer Sciences, 36*(2), 249-257.

Klecka, W. R. (1980). *Discriminant Analysis: Quantitative Applications in the Social Sciences*. Beverly Hills, CA: Sage.

Knoke, J. D. (1982). Discriminant Analysis with Discrete and Continuous Variables. *Biometrics, 38*(1), 191-200.

Kohavi, R. (1995). A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection. *Proceedings of the 14th International Joint Conference on Artificial Intelligence* (pp. 1137-1143). San Francisco, CA: Morgan Kaufmann.

Kolenikov, S. & Angeles, G. (2009). Socioeconomic Status Measurement With Discrete Proxy Variables: Is Principal Component Analysis A Reliable Answer? *Review of Income and Wealth, 55*(1), 128-165.

Kosambi, D. (1943). Statistics in Function Space. *Journal of Indian Mathematical Society, 7*, 76-88.

Kozma, L., Ilin, A. & Raiko, T. (2009). Binary Principal Component Analysis in the Netflix Collaborative Filtering Task. In *Proceedings of the IEEE International Workshop on Machine Learning for Signal Processing*, 1-4 September 2009, Grenoble, France (pp. 1-6). Retrieved from http://www.lkozma.net/mlsp09binary.pdf.

Kraha, A., Turner, H., Nimon, K., Zientek, L. R. & Henson, R. K. (2012). Tools to Support Interpreting Multiple Regression in the Face of Multicollinearity. *Frontiers in Psychology, 3,* 44-56.

Kriegel, H-P., Kröger, P. & Zimek, A. (2009). Clustering High-dimensional Data: A Survey on Subspace Clustering, Pattern-based Clustering and Correlation Clustering. *ACM Transactions on Knowledge Discovery from Data, 3*(1), 1-58.

Kristensen, P. (1992). Bias from Nondifferential but Dependent Misclassification of Exposure and Outcome. *Epidemiology, 3,* 210-215.

Krusińska, E. (1989). New Procedure for Selection of Variables in Location Model for Mixed Variable Discrimination. *Biometrics, 81*(5), 511-523.

Krzanowski, W. J. (1975). Discrimination and Classification using Both Binary and Continuous Variables. *Journal of the American Statistical Association, 70*(352), 782-790.

Krzanowski, W. J. (1977). The Performance of Fisher's Linear Discriminant Function under Non-optimal Conditions. *Technometrics, 19,* 191-200.

Krzanowski, W. J. (1979). Some Linear Transformations for Mixtures of Binary and Continuous Variables, with Particular Reference to Linear Discriminant Analysis. *Biometrika, 66*(1), 33-39.

Krzanowski, W. J. (1980). Mixtures of Continuous and Categorical Variables in Discriminant Analysis. *Biometrics, 36,* 493-499.

Krzanowski, W. J. (1982). Mixtures of Continuous and Categorical Variables in Discriminant Analysis: A Hypothesis Testing Approach. *Biometrics, 38*(4), 991-1002.

Krzanowski, W. J. (1983a). Stepwise Location Model Choice in Mixed-Variable Discrimination. *Applied Statistics, 32*(3), 260-266.

Krzanowski, W. J. (1983b). Distance between Populations using Mixed Continuous and Categorical Variables. *Biometrika, 70,* 235-243.

Krzanowski, W. J. (1984). On the Null Distribution of Distance between Two Groups, using Mixed Continuous and Categorical Variables. *Journal of Classification, 1*, 243-253.

Krzanowski, W. J. (1987). Selection of Variables to Preserve Multivariate Data Structure using Principal Components. *Applied Statistics, 36*(1), 22-33.

Krzanowski, W. J. (1993). The Location Model for Mixtures of Categorical and Continuous Variables. *Journal of Classification, 10*, 25-49.

Krzanowski, W. J. (1994). Quadratic Location Discriminant Functions for Mixed Categorical and Continuous Data. *Statistics & Probability Letters, 19*, 91-95.

Krzanowski, W. J. (1995). Selection of Variables, and Assessment of Their Performance, in Mixed-variable Discriminant Analysis. *Computational Statistics & Data Analysis, 19*, 419-431.

Krzanowski, W. J. (2000). *Principles of Multivariate Analysis: A User's Perspective*. New York, NY: Oxford University Press.

Krzanowski, W. J. (2003). Non-parametric Estimation of Distance between Groups. *Journal of Applied Statistics, 30*(7), 743-750.

Kshirsagar, A. M. (1972). *Multivariate Analysis*. New York, NY: Marcel Dekker, Inc.

Kshirsagar, A. M., Kocherlakota, S. & Kocherlakota, K. (1990). Classification Procedures using Principal Component Analysis and Stepwise Discriminant Function: Theory and Methods. *Communications in Statistics, 19*, 91-109.

Kullback, S. & Leibler, R. A. (1951). On Information and Sufficiency. *The Annals of Mathematical Statistics, 22*(1), 79-86.

Kumar, C. A. (2009). Analysis of Unsupervised Dimensionality Reduction Techniques. *Journal of Computer Science and Information Systems, 6*(2), 217-227.

Lachenbruch, P. A. (1967). An Almost Unbiased Method of Obtaining Confidents Intervals for the Probability of Misclassification in Discriminant Analysis. *Biometrics, 23*, 639-645.

Lachenbruch, P. A. (1975). *Discriminant Analysis*. New York, NY: Hafner Press.

Lachenbruch, P. A. & Mickey, M. R. (1968). Estimation of Error Rates in Discriminant Analysis. *Technometrics, 10*, 1-11.

Lachenbruch, P. A., Sneeringer, C. & Revo, L. T. (1973). Robustness of the Linear and Quadratic Discriminant Function to Certain Types of Non-normality. *Communications in Statistics, 1*, 39-56.

Lebart, L. (1975). L'orientation du Dépouillement de Certaines Enquêtes par l'analyse des Correspondences Multiples [The Orientation of the Analysis of Some Surveys by Multiple Correspondence Analysis]. *Consommation, 2*, 73-96.

Lebart, L., Morineau, A. & Tabard, N. (1977). *Techniques de la Description Statistique* [Statistical Description Techniques]. Paris: Dunod.

Lebart, L., Morineau, A. & Warwick, K. M. (1984). *Multivariate Descriptive Statistical Analysis: Correspondence Analysis and Related Techniques for Large Matrices*. New York, NY: John Wiley & Sons, Inc.

Lê Cao, K-A. & McLachlan, G. J. (2009). Statistical Analysis on Microarray Data: Selection of Gene Prognosis Signatures. In T. Pham (Eds.), *Computational Biology: Issues and Applications in Oncology* (pp. 55-76). London: Springer.

Leon, A. R., Soo, A. & Williamson, T. (2011). Classification with Discrete and Continuous Variables via General Mixed-Data Models. *Journal of Applied Statistics, 38*(5), 1021-1032.

LeRoux, B. & Rouanet, H. (2004). *Geometric Data Analysis: From Correspondence Analysis to Structured Data Analysis*. Dordrecht: Kluwer.

LeRoux, B. & Rouanet, H. (2010). *Multiple Correspondence Analysis: Quantitative Applications in the Social Sciences*. Thousand Oaks: Sage.

Lewis, D. P., Jebara, T. & Noble, W. S. (2006). Nonstationary Kernel Combination. In *Proceedings of the 23rd International Conference on Machine Learning*, 25-29 June 2006, Pittsburgh (pp. 553-560). New York, NY: ACM Press.

Li, D., Deogun, J. S. & Wang, K. (2007). Gene Function Classification Using Fuzzy *K*-Nearest Neighbor Approach. In *Proceedings of the IEEE International Conference on Granular Computing*, 2-4 November 2007, Fremont, CA (pp. 644-647). Washington, DC: IEEE Computer Society.

Li, D-C., Liu, C-W. & Hu, S. C. (2011). A Fuzzy-based Data Transformation for Feature Extraction to Increase Classification Performance with Small Medical Data Sets. *Artificial Intelligence in Medicine, 52*(1), 45-52.

Li, K. C. (1991). Sliced Inverse Regression for Dimension Reduction. *Journal of the American Statistical Association, 86*(414), 316-327.

Li, L. & Nachtsheim, C. J. (2007). Comment on "Fisher Lecture: Dimension Reduction in Regression". *Statistical Science, 22*(1), 36-39.

Li, Q. (2006). *An Integrated Framework of Feature Selection and Extraction for Appearance-based Recognition* (Unpublished doctoral dissertation). University of Delaware Newark, DE, USA.

Li, Y., Kittler, J. & Matas, J. (1999). Effective Implementation of Linear Discriminant Analysis for Face Recognition and Verification. *Proceedings of the 8th International Conference on Computer Analysis of Images and Patterns* (pp. 234-242). London: Springer-Verlag.

Liang, Y., Li, C., Gong, W. & Pan, Y. (2007). Uncorrelated Linear Discriminant Analysis based on Weighted Pairwise Fisher Criterion. *The Journal of the Pattern Recognition Society, 40*, 3606-3615.

Linting, M., Meulman, J. J., Groenen, P. J. & van der Kooij, A. J. (2007). Nonlinear Principal Components Analysis: Introduction and Application. *Psychological Methods, 12*(3), 336-358.

Liu, J. & Chen, S. (2006). Discriminant Common Vectors versus Neighbourhood Components Analysis and Laplacianfaces: A Comparative Study in Small Sample Size Problem. *Image and Vision Computing, 24*, 249-262.

LouisMarie, A. (2009). *Analysis of Multidimensional Poverty: Theory and Case Studies*. New York, NY: Springer-Verlag.

Lu, Y., Tian, Q., Sanchez, M., Neary, J., Liu, F. & Wang, Y. (2007). Learning Microarray Gene Expression Data by Hybrid Discriminant Analysis. *IEEE Multimedia Magazine, Special Issue on Multimedia Signal Processing and Systems in Health Care and Life Science, 14*(4), 22-31.

Lukibisi, F. B. & Lanyasunya, T. (2010). *Using Principal Component Analysis to Analyze Mineral Composition Data*. Paper Presented at the 12th Biennial KARI (Kenya Agricultural Research Institute) Scientific Conference on Socio Economics and Biometrics (pp. 1258-1268), 8-12 November, Nairobi, Kenya.

Lynn, H. S. & McCulloch, C. E. (2000). Using Principal Component Analysis and Correspondence Analysis for Estimation in Latent Variable Models. *Journal of the American Statistical Association, 95*(450), 561-572

Mahalanobis, P. C. (1936). On the Generalised Distance in Statistics. *Proceedings* of the *National Institute of Sciences of India, 2*(1), 49-55.

Mahat, N. I. (2006). *Some Investigations in Discriminant Analysis with Mixed Variables* (Unpublished doctoral dissertation). University of Exeter, London, UK.

Mahat, N. I., Krzanowski, W. J. & Hernandez, A. (2007). Variable Selection in Discriminant Analysis based on the Location Model for Mixed Variables. *Advances in Data Analysis and Classification, 1*(2), 105-122.

Mahat, N. I., Krzanowski, W. J. & Hernandez, A. (2009). Strategies for Non-Parametric Smoothing of the Location Model in Mixed-Variable Discriminant Analysis. *Modern Applied Science, 3*(1), 151-163.

Malin, B., Krueger, D. & Kubler, F. (2007). *Computing Stochastic Dynamic Economic Models with a Large Number of State Variables: A Description and Application of a Smolyak-Collocation Method* (NBER Working Paper No. NBER-WP-07-13517). Massachusetts, MA: National Bureau of Economic Research. Retrieved from http://www.nber.org/papers/w13517.

Malinvaud, E. (1987). *Data Analysis in Applied Socio-economic Statistics with Special Consideration of Correspondence Analysis*. Paper Presented at the Marketing Science Conference 1987, Jouy en Josas, France.

Mardia, K. V., Kent, J. T. & Bibby, J. M. (1979). *Multivariate Analysis*. London: Academic Press, Inc.

Marks, S. & Dunn, O. J. (1974). Discriminant Function when Covariance Matrices are Unequal. *Journal of the American Statistical Association, 69*, 555-559.

Martinez, A. M. & Kak, A. C. (2001). PCA versus LDA. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 23*(2), 228-233.

Massey, W. F. (1965). Principal Components Regression in Exploratory Statistical Research *Journal of American Statistical Association, 60*, 234-246.

Matusita, K. (1956). Decision Rule, based on the Distance, for the Classification Problem. *Annals of the Institute of Statistical Mathematics, 16*, 305-315.

Mažgut, J., Tiňo, P., Bodén, M. & Yan, H. (2010). Multilinear Decomposition and Topographic Mapping of Binary Tensors. In K. Diamantaras, W. Duch & L. S. Iliadis (Eds.), *Artificial Neural Networks: Proceedings of the 20th International Conference on Artificial Neural Networks - Part I*, 15-18 September 2010, Thessaloniki, Greece (pp. 317-326). Berlin, Heidelberg: Springer-Verlag.

McCabe, G. P. (1984). Principal Variables. *Technometrics, 26*(2), 137-144.

McCulloch, W. S. & Pitts, W. (1943). A Logical Calculus of the Ideas Immanent in Nervous Activity. *Bulletin of Mathematical Biophysics, 5*(4), 115-133.

McGarigal, K., Cushman, S. A. & Stafford, S. G. (2000). *Multivariate Statistics for Wildlife and Ecology Research*. New York, NY: Springer-Verlag.

McKay, R. J. & Campbell, N. A. (1982). Variable Selection Techniques in Discriminant Analysis II: Allocation. *British Journal of Mathematical and Statistical Psychology, 35*, 30-41.

McLachlan, G. J. (1976). A Criterion for Selecting Variables for the Linear Discriminant Function. *Biometrics, 32*(3), 529-534.

McLachlan, G. J. (1992). *Discriminant Analysis and Statistical Pattern Recognition*. New York, NY: John Wiley & Sons, Inc.

McLachlan, G. J. (2004). *Discriminant Analysis and Statistical Pattern Recognition* (2nd ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Merbouha, A. & Mkhadri, A. (2004). Regularization of the Location Model in Discrimination with Mixed Discrete and Continuous Variables. *Computational Statistics and Data Analysis, 45*, 563-576.

Messaoud, R. B., Boussaid, O. & Rabaséda, S. L. (2007). A Multiple Correspondence Analysis to Organize Data Cubes. In O. Vasilecas, J. Eder & A. Caplinskas (Eds.), *Databases and Information Systems IV: Frontiers in Artificial Intelligence and Applications* (pp. 133-146). Amsterdam: IOS Press.

Meulman, J. J., van Der Kooij, A. J. & Heiser, W. J. (2004). Principal Components Analysis with Nonlinear Optimal Scaling Transformations for Ordinal and Nominal Data. In D. Kaplan (Eds.), *The SAGE Handbook of Quantitative Methodology for the Social Sciences* (pp. 49-70). Thousand Oaks: Sage.

Mitchell, T. M. (1997). *Machine Learning*. New York, NY: McGraw-Hill.

Moeinzadeh, H., Mohammadi, M-M., Akbari, A. & Nasersharif, B. (2009). Robust Speech Recognition using Evolutionary Class-dependent LDA. In *Proceedings of the 11th Annual Conference Companion on Genetic and Evolutionary Computation*, 8-12 July 2009, Montréal Québec, Canada (pp. 2109-2114). New York, NY: ACM Press.

Mood, A. M., Graybill, F. A. & Boes, D. C. (1973). *Introduction to the Theory of Statistics* (3rd ed.). New York, NY: McGraw-Hill.

Moore, D. H. (1973). Evaluation of Five Discriminant Procedures for Binary Variables. *Journal of the American Statistical Association, 68*, 399-404.

Moravec, P. (2005). Testing Dimension Reduction Methods for Text Retrieval. In K. Richta, V. Snášel & J. Pokorný (Eds.), *Proceedings of the Dateso '05 Annual International Workshop on Databases, Texts, Specifications and Objects* (pp. 113-124), 13-15 April, Desna, Czech Republic. Retrieved from http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS//Vol-129/paper15.pdf.

Morrison, D. F. (1976). *Multivariate Statistical Methods* (2nd ed.). New York, NY: McGraw-Hill.

Moussa, M. A. (1980). Discrimination and Allocation using A Mixture of Discrete and Continuous Variables with Some Empty States. *Computer Programs in Biomedicine, 12*(2-3), 161-171.

Murray, G. D. (1977). A Cautionary Note on Selection of Variables in Discriminant Analysis. *Journal of the Royal Statistical Society. Series C: Applied Statistics, 26*(3), 246-250.

Muthén, B. (1978). Contribution to Factor Analysis of Dichotomized Variables. *Psychometrika, 43*(4), 551-560.

Nedal Omar, M. A. A. (2010). *Detection of Multicollinearity in Multiple Linear Regression* (Unpublished master's thesis). Al-Azhar University, Gaza, Palestine.

Negahban, S. & Wainwright, M. J. (2011). Estimation of (near) Low-rank Matrices with Noise and High-dimensional Scaling. *The Annals of Statistics, 39*(2), 1069-1097.

Nenadić, O. & Greenacre, M. J. (2007). Correspondence Analysis in R, with Two- and Three-dimensional Graphics: The ca Package. *Journal of Statistical Software, 20*(3), 1-13.

Nenadic, Z. (2007). Information Discriminant Analysis: Feature Extraction with an Information-Theoretic Objective. *IEEE Transactions on Pattern Analysis and Machine Intelligent, 29*(8), 1394-1407.

Nguyen, D. V. & Rocke, D. M. (2002a). Tumor Classification by Partial Least Squares using Microarray Gene Expression Data. *Bioinformatics, 18*(1), 39-50.

Nguyen, D. V. & Rocke, D. M. (2002b). Multi-class Cancer Classification via Partial Least Squares with Gene Expression Profiles. *Bioinformatics, 18*(9), 1216-1226.

Nie, F., Xiang, S., Song, Y. & Zhang, C. (2007). Extracting the Optimal Dimensionality for Discriminant Analysis. *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing* (pp. 617-620). Retrieved from http://ieeexplore.ieee.org/xpls/abs.

Nishisato, S. (1980). *Analysis of Categorical Data: Dual Scaling and its Applications*. Toronto: University of Toronto Press.

Njong, A. M. & Ningaye, P. (2008). *Characterizing Weights in the Measurement of Multidimensional Poverty: An Application of Data-driven Approaches to*

Cameroonian Data (OPHI Working Paper No. OPHI-WP-08-21). Oxford: Oxford Poverty & Human Development Initiative. Retrieved from http://www.ophi.org.uk/wp-content/uploads/OPHI-wp21.pdf.

Oja, E. (1992). Principal Components, Minor Components, and Linear Neural Networks. *Neural Networks, 5*, 927-935.

Okpeku, M., Yakubu, A., Peters, S. O., Ozoje, M. O., Ikeobi, C. O. N., Adebambo, O. A. & Imumorin, I. G. (2011). Application of Multivariate Principal Component Analysis to Morphological Characterization of Indigenous Goats in Southern Nigeria. *Acta Agriculturae Slovenica, 98*(2), 101-109.

Olkin, I. & Tate, R. F. (1961). Multivariate Correlation Models with Mixed Discrete and Continuous Variables. *The Annals of Mathematical Statistics, 32*(2), 448-465.

Osborne, M. R. (1976). On the Computation of Stepwise Regressions. *Australia Computer Journal, 8*, 61-68.

Pasiouras, F., Tanna, S. & Zopounidis, C. (2005). *Application of Quantitative Techniques for the Prediction of Bank Acquisition Targets: Computers and Operations Research* (Vol. 5). New Jersey: World Scientific.

Pavlenko, T. (2001). *Feature Informativeness, Curse-of-Dimensionality and Error Probability in Discriminant Analysis* (Doctoral dissertation, Lund University). Retrieved from http://www.maths.lth.se/matstat/publications/phdtheses.

Pavlenko, T. & von Rosen, D. (2002). Bayesian Network Classifiers in a High Dimensional Framework. *Proceedings of the 18th Annual Conference on Uncertainty in Artificial Intelligence* (pp. 397-404). San Francisco, CA: Morgan Kaufmann.

Pearson, K. (1901). On Lines and Planes of Closest Fit to Systems of Points in Space. *Philosophical Magazine, 6*(2), 559-572.

Pechenizkiy, M. (2005). The Impact of Feature Extraction on the Performance of a Classifier: kNN, Naïve Bayes and C4.5. In B. Kégl & G. Lapalme (Eds.), *Proceedings of the 18th Canadian Society Conference on Advances in Artificial Intelligence*, 9-11 May 2005, Victoria, Canada (pp. 268-279). Berlin, Heidelberg: Springer-Verlag.

Penchev, P. N., Argirov, O. K. & Andreev, G. N. (1994). Mass Spectra Classification according to Substructures and Molecular Formula using Artificial Neural Networks. *Analytical Laboratory, 3,* 29-33.

Peter, G. M., Joop, T. & Charles, O. (1997). Multiple Correspondence Analysis as A Tool for Quantification or Classification of Career Data. *Journal of Educational and Behavioral Statistics, 22*(4), 447-477.

Pfurtscheller, G., Neuper, Ch., Flotzinger, D. & Pregenzer, M. (1997). EEG-based Discrimination between Imagination of Right and Left Hand Movement. *Electroencephalography and Clinical Neurophysiology, 103*(6), 642-651.

Ping, H. (2005). *Classification Methods and Applications to Mass Spectral Data* (Unpublished doctoral dissertation). Hong Kong Baptist University, Hong Kong.

Pinkowski, B. (1987). Discrete Discriminant Models: A Performance Simulation with Reference to Expert Systems' Applications. In R. A. Gagliano (Eds.), *Proceedings of the 20th Annual Symposium on Simulation*, Florida (pp. 103-119). Los Alamitos, CA: IEEE Computer Society Press.

Pinto, L. F. B., Packer, I. U., De Melo, C. M. R., Ledur, M. C. & Coutinho, L. L. (2006). Principal Components Analysis Applied to Performance and Carcass Traits in The Chicken. *Animal Research, 55*(5), 419-425.

Prats-Moltalbán, J. M., Ferrer, A., Malo, J. L. & Gorbeña, J. (2006). A Comparison of Different Discriminant Analysis Techniques in a Steel Industry Welding Process. *Chemometrics and Intelligent Laboratory Systems, 80,* 109-119.

Preacher, K. J. & MacCallum, R. C. (2002). Exploratory Factor Analysis in Behavior Genetics Research: Factor Recovery with Small Sample Sizes. *Behavior Genetics, 32,* 153-161.

Press, S. J. & Wilson, S. (1978). Choosing between Logistic Regression and Discriminant Analysis. *Journal of the American Statistical Association, 73,* 699-705.

Qiao, Z., Zhou, L. & Huang, J. Z. (2008). Effective Linear Discriminant Analysis for High Dimensional, Low Sample Size Data. In S. I. Ao, L. Gelman, D. W. L. Hukins, A. Hunter & A. M. Korsunsky (Eds.), *Proceedings of the World*

Congress on Engineering, 2, 2-4 July 2008, London (pp. 1070-1075). Newswood Limited.

Qu, Y., Adam, B. L., Thornquist, M., Potter, J. D., Thompson, M. L., Yasui, Y., Davis, J., Schellhammer, P. F., Cazares, L., Clements, M., Jr. Wright, G. L. & Feng, Z. (2003). Data Reduction using A Discrete Wavelet Transform in Discriminant Analysis of Very High Dimensionality Data. *Biometrics, 59*(1), 143-151.

Quinlan, J. R. (1986). Induction of Decision Trees. *Machine Learning, 1*(1), 81-106.

Raiko, T., Ilin, A. & Karhunen, J. (2007). Principal Component Analysis for Large Scale Problems with Lots of Missing Values. In J. N. Kok, J. Koronacki, R. L. de Mantaras (Eds.), *Proceedings of the 18th European Conference on Machine Learning*, September 2007, Warsaw, Poland (pp. 691-698). Berlin, Heidelberg: Springer-Verlag.

Rao, C. R. (1952). *Advanced Statistical Methods in Multivariate Analysis*. New York, NY: John Wiley & Sons, Inc.

Rao, C. R. (1964). The Use and Interpretation of Principal Component Analysis in Applied Research. *Sankhyā: The Indian Journal of Statistics, 26*(4), 329-358.

Raudys, S. & Duin, R. P. W. (1998). Expected Classification Error of the Fisher Linear Classifier with Pseudo-Inverse Covariance Matrix. *Pattern Recognition Letters, 19*(5-6), 385-392.

Reise, S. P., Waller, N. G. & Comrey, A. L. (2000). Factor Analysis and Scale Revision. *Psychological Assessment, 12*, 287-297.

Rencher, A. C. (1998). *Multivariate Statistical Inference and Applications*. New York, NY: John Wiley & Sons, Inc.

Rencher, A. C. (2002). *Methods of Multivariate Analysis: Wiley Series in Probability and Statistics* (2nd ed.). New York, NY: John Wiley & Sons, Inc.

Ripley, B. D. (1996). *Pattern Recognition and Neural Networks*. Cambridge: Cambridge University Press.

Robinson, L. D. & Jewell, N. P. (1991). Some Surprising Results about Covariate Adjustment in Logistic Regression. *International Statistical Review, 58*, 227-240.

Rust, J. (1997). Using Randomization to Break the Curse of Dimensionality. *Econometrica, 65*(3), 487-516.

Sahn, D. & Stifel, D. (2003). Exploring Alternative Measures of Welfare in the Absence of Expenditure Data. *Review of Income and Wealth, 49*, 463-489.

San, O. M., Huynh, V-n. & Nakamori, Y. (2004). An Alternative Extension of the *k*-Means Algorithm for Clustering Categorical Data. *International Journal of Applied Mathematics and Computer Science, 14*(2), 241-247.

Saporta, G. (1990). Simultaneous Analysis of Qualitative and Quantitative Data. In *Proceedings of the 35th Scientific Meeting of the Italian Statistical Society, 1*, 18-21 April 1990, Padova (pp. 63-72). Retrieved from http://cedric.cnam.fr/~saporta/SAQQD.pdf.

Saporta, G. & Tambrea, N. (1993). About the Selection of the Number of Components in Correspondence Analysis. In J. Janssen & C. H. Skiadas (Eds.), *Applied Stochastic Models and Data Analysis* (pp. 846-856). Singapore: World Scientific Publishing.

Schäfer, J. & Strimmer, K. (2005). A Shrinkage Approach to Large-scale Covariance Matrix Estimation and Applications for Functional Genomics. *Statistical Applications in Genetics and Molecular Biology, 4*(1), 176-208.

Schein, A. I., Saul, L. K. & Ungar, L. H. (2003). A Generalized Linear Model for Principal Component Analysis of Binary Data. In C. M. Bishop & B. J. Frey (Eds.), *Proceedings of the 9th International Workshop on Artificial Intelligence and Statistics* (pp. 14-21), 3-6 January, Key West, FL.

Schmitz, P. I. M., Habbema, J. D. F. & Hermans, J. (1983). The Performance of Logistic Discrimination on Myocardial Infarction Data, in Comparison with Some Other Discriminant Analysis Methods. *Statistics in Medicine, 2*, 199-205.

Schumacher, M., Roßner, R. & Vach, W. (1996). Neural Networks and Logistic Regression: Part I. *Computational Statistics & Data Analysis, 21*, 661-682.

Scott, D. W. & Factor, L. E. (1981). Monte Carlo Study of Three Data-based Nonparametric Probability Density Estimators. *Journal of the American Statistical Association, 76*, 9-15.

Sebban, M., Nock, R. & Lallich, S. (2002). Stopping Criterion for Boosting-based Data Reduction Techniques: from Binary to Multiclass Problems. *Journal of Machine Learning Research, 3*, 863-885.

Seber, G. A. F. (1984). *Multivariate Observations*. New York, NY: John Wiley & Sons, Inc.

Shin, K-S., Lee, T. S. & Kim, H-j. (2005). An Application of Support Vector Machines in Bankruptcy Prediction Model. *Expert Systems with Applications, 28*(1), 127-135.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. New York, NY: Chapman & Hall/CRC.

Simon, R., Radmacher, M. D., Dobbin, K. & McShane, L. M. (2003). Pitfalls in the Use of DNA Microarray Data for Diagnostic and Prognostic Classification. *Journal of the National Cancer Institute, 95*(1), 14-18.

Sirovich, L. & Kirby, M. (1987). Low-dimensional Procedure for the Characterization of Human Faces. *Journal of the Optical Society America, 4*(3), 519-524.

Sithole, M. M. (1992). *Variable Selection in Principal Component Analysis: Using Measures of Multivariate Association* (Master's thesis, Curtin University of Technology). Retrieved from http://espace.library.curtin.edu.au:80/R?func=dbin-jump-full&local_base=gen01-era02&object_id=11165.

Skurichina, M. & Duin, R. P. W. (1996). Stabilizing Classifiers for Very Small Sample Sizes. In *Proceedings of the 13th International Conference on Pattern Recognition, 2*, 25-29 August 1996, Vienna, Austria (pp. 891-896). Washington, DC: IEEE Computer Society Press.

Smith, C. A. B. (1947). Some Examples of Discrimination. *Annals of Eugenics, 13*, 272-283.

Snapinn, S. M. & Knoke, J. D. (1985). An Evaluation of Smoothed Classification Error-rate Estimators. *Technometrics, 27*, 199-206.

Snapinn, S. M. & Knoke, J. D. (1989). Estimation of Error Rates in Discriminant Analysis with Selection of Variables. *Biometrics, 45*, 289-299.

Song, Y., Huang, J., Zhou, D., Zha, H. & Giles, C. L. (2007). IKNN: Informative K-Nearest Neighbor Pattern Classification. In J. N. Kok, J. Koronacki, R. L. de Mantaras & D. Mladenic (Eds.), *Proceedings of the 11th European Conference on Principles and Practice of Knowledge Discovery in Databases*, 17-21 September 2007, Warsaw, Poland (pp. 248-264). Berlin, Heidelberg: Springer-Verlag.

Sourial, N., Wolfson, C., Zhu, B., Quail, J., Fletcher, J., Karunananthan, S., Bandeen-Roche, K., Béland, F. & Bergman, H. (2010). Correspondence Analysis is a Useful Tool to Uncover the Relationships among Categorical Variables. *Journal of Clinical Epidemiology, 63*(6), 638-646.

Speed, T. (2003). *Statistical Analysis of Gene Expression Microarray Data*. Boca Raton: Chapman & Hall/CRC.

Stephens, M. A. (1974). EDF Statistics for Goodness of Fit and Some Comparisons. *Journal of the American Statistical Association, 69*(347), 730-737.

Stern, H.S. (1996). Neural Networks in Applied Statistics. *Technometrics, 38*, 217-225.

Stevens, J. (1986). *Applied Multivariate Statistics for the Social Sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Swets, D. L. & Weng, J. J. (1996). Using Discriminant Eigenfeatures for Image Retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 18*(8), 831-836.

Tatti, N., Mielikäinen, T., Gionis, A. & Mannila, H. (2006). What is the Dimension of Your Binary Data? In *Proceedings of the Sixth International Conference on Data Mining*, 18-22 December 2006, Hong Kong (pp. 603-612). Washington, DC: IEEE Computer Society Press.

Tay, F. E. H. & Cao, L. (2001). Application of Support Vector Machines in Financial Time Series Forecasting. *The International Journal of Management Science, 29*(4), 309-317.

Tay, F. E. H. & Cao, L. (2002). Modified Support Vector Machines in Financial Time Series Forecasting. *Neurocomputing, 48*(1-4), 847-861.

Tenenhaus, M. & Young, F. W. (1985). An Analysis and Synthesis of Multiple Correspondence Analysis, Optimal Scaling, Dual Scaling, Homogeneity Analysis and Other Methods for Quantifying Categorical Multivariate Data. *Psychometrika, 50*(1), 91-119.

Thomaz, C. E., Gillies, D. F. & Feitosa, R. Q. (2001). Small Sample Problem in Bayes Plug-in Classifier for Image Recognition. In *Proceedings of International Conference on Image and Vision Computing*, New Zealand (pp. 295-300). Dunedin: Wickliffe Limited.

Thomaz, C. E., Kitani, E. C. & Gillies, D. F. (2006). A Maximum Uncertainty LDA-based Approach for Limited Sample Size Problems - with Application to Face Recognition. *Journal of the Brazilian Computer Society, 12*(2), 7-18.

Thompson, B. (2006). *Foundations of Behavioral Statistics: An Insight-Based Approach*. New York, NY: Guilford Press.

Tian, Y., Guo, P. & Lyu, M. R. (2005). Comparative Studies on Feature Extraction Methods for Multispectral Remote Sensing Image Classification. In *Proceedings of the International Conference on Systems, Man and Cybernetics*, 10-12 October 2005, Waikoloa, Hawaii (pp. 1275-1279). Berlin, Heidelberg: Springer-Verlag.

Tu, J. V. (1996). Advantages and Disadvantages of using Artificial Neural Networks versus Logistic Regression for Predicting Medical Outcomes. *Journal of Clinical Epidemiology, 49*(11), 1225-1231.

Turk, M. A. & Pentland, A. P. (1991). Eigenfaces for Recognition. *Journal of Cognitive Neuroscience, 3*(1), 71-86.

Urbakh, V. Y. (1971). Linear Discriminant Analysis: Loss of Discriminating Power when a Variate is Omitted. *Biometrics, 27*(3), 531-534.

Ushakov, N. G. (2001). Density of a Probability Distribution. In M. Hazewinkel (Eds.), *Encyclopedia of Mathematics*. New York, NY: Springer-Verlag.

Utsugi, A. (1997). Topology Selection for Self-organizing Maps. *Network: Computation in Neural Systems, 7*, 727-740.

Valette-Florence, P. & Rapacchi, B. (1991). Improvements in Means-End Chain Analysis using Graph Theory and Correspondence Analysis. *Journal of Advertising Research*, 30-45.

van Buuren, S. & de Leeuw, J. (1992). Equality Constraints in Multiple Correspondence Analysis. *Multivariate Behavioral Research, 27*(4), 567-583.

van der Heijden, P. G. M., de Falguerolles, A. & de Leeuw, J. (1989). A Combined Approach to Contingency Table Analysis using Correspondence Analysis and Log-linear Analysis. *Applied Statistics, 38,* 249-292.

van der Heijden, P. G. M., Mooijaart, Ab. & Takane, Y. (1994). Correspondence Analysis and Contingency Table Models. In J. Blasius & M. J. Greenacre (Eds.), *Correspondence Analysis in the Social Sciences: Recent Developments and Applications* (pp. 79-111). London: Academic Press.

Vapnik, V. (1979). *Estimation of Dependences Based on Empirical Data*. New York, NY: Springer-Verlag.

Vapnik, V. (1995). *The Nature of Statistical Learning Theory*. New York, NY: Springer-Verlag.

Vapnik, V., Golowich, S. & Smola, A. (1997). Support Vector Method for Function Approximation, Regression Estimation and Signal Processing. In M. Mozer, M. Jordan & T. Petsche (Eds.), *Advances in Neural Information Processing Systems* (pp. 281-287). Cambridge, MA: MIT Press.

Velicer, W. F. & Fava, J. L. (1998). Effects of Variable and Subject Sampling on Factor Pattern Recovery. *Psychological Methods, 3*, 231-251.

Victor, B. L. & Zhang, G. P. (1999). The Effect of Misclassification Costs on Neural Network Classifiers. *Decision Science, 30,* 659-682.

Vinay, V., Cox, I. J., Wood, K. & Milic-Frayling, N. (2005). A Comparison of Dimensionality Reduction Techniques for Text Retrieval. In *Proceedings of 4th International Conference on Machine Learning and Applications,* 15-17 December 2005 (pp. 293-298). Washington, DC: IEEE Computer Society Press.

Vlachonikolis, I. G. & Marriott, F. H. C. (1982). Discrimination with Mixed Binary and Continuous Data. *Applied Statistics, 31*(1), 23-31.

Vyas, S. & Kumaranayake, L. (2006). Constructing Socio-economic Status Indices: How to Use Principal Components Analysis. *Health Policy and Planning, 21*(6), 459-468.

Wahl, P. W. & Kronmal, R. A. (1977). Discriminant Functions when Covariances are Unequal and Sample Sizes are Moderate. *Biometrics, 33,* 479-484.

Wakaki, H. (1990). Comparison of Linear and Quadratic Discriminant Functions. *Biometrika, 77,* 227-229.

Wang, X. & Kabán, A. (2005). Finding Uninformative Features in Binary Data. In M. Gallagher, J. Hogan & F. Maire (Eds.), *Proceedings of the 6th International Conference on Intelligent Data Engineering and Automated Learning,* July 2005, Brisbane, Australia (pp. 40-47). Berlin, Heidelberg: Springer-Verlag.

Wang, X. & Tang, X. (2004). Experimental Study on Multiple LDA Classifier Combination for High Dimensional Data Classification. In F. Roli, J. Kittler & T. Windeatt (Eds.), *Proceedings of the 5th International Workshop on Multiple Classifier Systems*, 9-11 June 2004, Cagliari, Italy (pp. 344-353). Berlin, Heidelberg: Springer-Verlag.

Warmuth, M. K. & Kuzmin, D. (2008). Randomized Online PCA Algorithms with Regret Bounds that are Logarithmic in the Dimension. *Journal of Machine Learning Research, 9*(10), 2287-2320.

Wasserman, L. (2006). *All of Nonparametric Statistics*. New York, NY: Springer-Verlag.

Webb, L. (2002). Deliberate Self-harm in Adolescence: A Systematic Review of Psychological and Psychosocial Factors. *Journal of Advanced Nursing, 38*(3), 235-244.

Weiner, J. M. & Dunn, O. J. (1966). Elimination of Variate in Linear Discrimination Problems. *Biometrics, 22*(2), 268-275.

Wernecke, K. D. (1992). A Coupling Procedure for the Discrimination of Mixed Data. *Biometrics, 48*(2), 497-506.

West, M., Blanchette, C., Dressman, H., Huang, E., Ishida, S., Spang, R., Zuzan, H., Jr. Olson, J. A., Marks, J. R. & Nevins, J. R. (2001). Predicting the Clinical Status of Human Breast Cancer using Gene Expression Profiles. In P. J. Bickel (Eds.), *Proceedings of the National Academy of Sciences of the United States of America*, *98*(20), 11462-11467.

Wilbur, J. D., Ghosh, J. K., Nakatsu, C. H., Brouder, S. M. & Doerge, R. W. (2002). Variable Selection in High-Dimensional Multivariate Binary Data with Application to the Analysis of Microbial Community DNA Fingerprints. *Biometrics, 58*(2), 378-386.

Wilk, M. B. & Gnanadesikan, R. (1968). Probability Plotting Methods for the Analysis of Data. *Biometrika, 55*(1), 1-17.

Williams, B. K. & Titus, K. (1988). Assessment of Sampling Stability in Ecological Applications of Discriminant Analysis. *Ecology, 69*(4), 1275-1285.

Wold, H. (1966). Nonlinear Estimation by Iterative Least Squares Procedures. In F. N. David (Eds.), *Research Papers in Statistics* (pp. 411-444). New York, NY: John Wiley & Sons, Inc.

Wold, H. (1973). Nonlinear Iterative Partial Least Squares (NIPALS) Modeling: Some Recent Developments. In P. Krishnaiah (Eds.), *Multivariate Analysis III* (pp. 383-407). New York, NY: Academic Press.

Wold, S., Esbensen, K. & Geladi, P. (1987). Principal Component Analysis. C*hemometrics Intelligent Laboratory Systems, 2*, 37-52.

Wood, M., Jolliffe, I. T. & Horgan, G. W. (2005). Variable Selection for Discriminant Analysis of Fish Sounds using Matrix Correlations. *Journal of Agricultural, Biological and Environmental Statistics, 10*(3), 321-336.

Wood, S. N. (2006). *Generalized Additive Models: An Introduction with R*. FL: Chapman & Hall/CRC.

Worth, A. P. & Cronin, M. T. D. (2003). The Use of Discriminant Analysis, Logistic Regression and Classification Tree Analysis in the Development of Classification Models for Human Health Effects. *Journal of Molecular Structure: Theochem, 622*(1-2), 97-111.

Wu, Y., Tian, Q. & Huang, T. S. (2000). Discriminant-EM Algorithm with Application to Image Retrieval. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 13-15 June 2000, South Carolina (pp. 222-227). Retrieved from http://vision.eecs.northwestern.edu/papers/conference/2000/cvpr00a.pdf.

Xing, E. P., Jordan, M. I. & Karp, R. M. (2001). Feature Selection for High-Dimensional Genomic Microarray Data. *Proceedings of the 18th International Conference on Machine Learning* (pp. 601-608). San Francisco, CA: Morgan Kaufmann.

Xu, L., Krzyżak, A. & Suen, C. Y. (1992). Methods of Combining Multiple Classifiers and Their Applications to Handwriting Recognition. *IEEE Transactions on Systems, Man and Cybernetics, 22*(3), 418-435.

Xu, P., Brock, G. N. & Parrish, R. S. (2009). Modified Linear Discriminant Analysis Approaches for Classification of High-Dimensional Microarray Data. *Computational Statistics and Data Analysis, 53*(5), 1674-1687.

Yakubu, A. & Ayoade, J. A. (2009). Application of Principal Component Factor Analysis in Quantifying Size and Morphological Indices of Domestic Rabbits. *International Journal of Morphology, 27*, 1013-1017.

Yamada, F. & Nishisato, S. (1993). Several Mathematical Properties of Dual Scaling as Applied to Dichotomous Item Category Data. *Japanese Journal of Behaviormetrics, 29*(1), 56-63.

Yang, C., Peng, H. & Wang, J. (2008). A New Feature Extraction Approach Based on Sentence Element Analysis. In *Proceedings of the International Conference on Computational Intelligence and Security, 1*, 13-17 December 2008 (pp. 90-95). Washington, DC: IEEE Computer Society Press.

Yang, J. & Yang, J. Y. (2001). Optimal FLD Algorithm for Facial Feature Extraction. In D. P. Casasent & E. L. Hall (Eds.), *SPIE Proceedings of the Intelligent Robots and Computer Vision XX: Algorithms, Techniques and*

Active Vision, 4572 (pp. 438-444). Retrieved from http://dx.doi.org/10.1117/12.444212.

Yang, J. & Yang, J. Y. (2003). Why Can LDA be Performed in PCA Transformed Space? Rapid and Brief Communication. *Pattern Recognition, 36*(2), 563-566.

Yang, J., Ye, H. & David, Z. (2004). A New LDA-KL Combined Method for Feature Extraction and Its Generalization. *Pattern Analysis and Applications, 7*(2), 225-235.

Yang, K. & Trewn, J. (2004). *Multivariate Statistical Methods in Quality Management*. New York, NY: McGraw-Hill.

Ye, J., Janardan, R. & Li, Q. (2004). GPCA: An Efficient Dimension Reduction Scheme for Image Compression and Retrieval. In *Proceedings of the 10th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 22-25 August 2004, Washington (pp. 354-363). New York, NY: ACM Press.

Ye, J. & Wang, T. (2006). Regularized Discriminant Analysis for High Dimensional, Low Sample Size Data. In *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 20-23 August 2006, Pennsylvania (pp. 454-463). New York, NY: ACM Press.

Ye, J. & Xiong, T. (2006). Computational and Theoretical Analysis of Null Space and Orthogonal Linear Discriminant Analysis. *Journal of Machine Learning Research, 7*, 1183-1204.

Ye, J., Xiong, T., Li, Q., Janardan, R., Bi, J., Cherkassky, V. & Kambhamettu, C. (2006). Efficient Model Selection for Regularized Linear Discriminant Analysis. In *Proceedings of the 15th ACM International Conference on Information and Knowledge Management*, 5-11 November 2006, Virginia (pp. 532-539). New York, NY: ACM Press.

Young, F. W., Takane, Y. & de Leeuw, J. (1978). The Principal Components of Mixed Measurement Level Multivariate Data: An Alternating Least Squares Method with Optimal Scaling Features. *Psychometrika, 43*(2), 279-281.

Young, P. D. (2009). *Dimension Reduction and Missing Data in Statistical Discrimination* (Doctoral dissertation, USA Baylor University). Retrieved from https://beardocs.baylor.edu/xmlui/bitstream/handle/2104/5543/philip_young_phd.pdf.

Yu, H. & Yang, J. (2001). A Direct LDA Algorithm for High-Dimensional Data with Application to Face Recognition. *Pattern Recognition, 34*(10), 2067-2070.

Yu, L. (2009). Feature Cluster Selection for High-Throughput Data Analysis. *International Journal of Data Mining and Bioinformatics, 3*(2), 177-191.

Zabaleta, L., Barron, L. J. R., Pérez-Elortondo, F. J., Virto, M., Nájera, A. I., Agirre, E., de Gordoa, J. C. R., Amores, G., Menéndez, S., de Renobales, M. & Albisu, M. (2012). *Multiple Correspondence Analysis to Uncover Associations among Cheese's Sensory Variables*. Postal Presented at the International Conference of Sensometrics, 10-13 July, Rennes, France.

Zass, R. & Shashua, A. (2007). Nonnegative Sparse PCA. *Proceedings of the Twentieth Annual Conference on Advances in Neural Information Processing Systems, 19* (pp. 1561-1568). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.72.7798.

Zhang, X. & Jia, Y. (2007). A Linear Discriminant Analysis Framework based on Random Subspace for Face Recognition. *Journal of the Pattern Recognition Society, 40*, 2585-2591.

Zhao, W., Chellappa, R. & Nandhakumar, N. (1998). Empirical Performance Analysis of Linear Discriminant Classifiers. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, June 1998, Santa Barbara, CA (pp. 164-169). Washington, DC: IEEE Computer Society Press.

Zhao, W., Chellappa, R., Phillips, P. J. & Rosenfeld, A. (2003). Face Recognition: A Literature Survey. *ACM Computing Surveys, 35*(4), 399-458.

Zhu, M. (2001). *Feature Extraction and Dimension Reduction with Applications to Classification and the Analysis of Co-occurrence Data* (Doctoral dissertation, Stanford University). Retrieved from http://www.stanford.edu/~hastie/THESES/mu_zhu.pdf.

Zhu, M. & Martinez, A. M. (2006). Subclass Discriminant Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 28*(8), 1274-1286.

Zivkovic, Z. & Verbeek, J. J. (2006). Transformation Invariant Component Analysis for Binary Images. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 254-259). Retrieved from http://lear.inrialpes.fr/pubs/2006/ZV06.

Zuo, W., Zhang, D., Yang, J. & Wang, K. (2006). BDPCA plus LDA: A Novel Fast Feature Extraction Technique for Face Recognition. *IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics, 36*(4), 946-953.