FUZZY DATA ANALYSIS

Sato-Ilic, Mika
University of Tsukuba, Tsukuba, Ibarali, Japan

Keywords: fuzzy clustering, classification structure, HDLSS data, symbolic data, similarity, dissimilarity, generalized aggregation operator, nonlinear fuzzy clustering model

Contents

Summary

The amount of data is growing at an exponential rate. We are faced with a challenge to analyze, process and extract useful information from the vast amount of data. Traditional data analysis techniques have contributed immensely in the area of data analysis but we believe that the fuzzy logic can be complementary and can process complicated data sets. This chapter provides two kinds of novel fuzzy data analyses.

1. Introduction

Fuzzy logic reflects the pervasiveness of imprecision and uncertainty which exists in the real world. On the other hand, hard computing does not reflect this imprecision and uncertainty. The guiding principle of fuzzy logic is to exploit the tolerance of imprecision and uncertainty in order to achieve tractability, robustness, and low solution cost.

Recently, in the area of data analysis, many new methods have been proposed. One reason for this is that traditional data analysis does not adequately reflect the imprecision and uncertainty of real world data. For instance, analysis methods for uncertainty data including interval-valued data, fuzzy data, modal data, functional data, and categorical data have been proposed. Data which have a much larger number of variables than the number of objects are also important concerns in multivariable data analysis. Such data have been obtained in various areas, such as genomics, bioinformatics, chemometrics, brain sciences, and functional data analysis. Conventional multivariate analyses cannot treat such data. Huge amounts of data (big data) are also a problem and data mining is placing an increased emphasis on revealing the latent structure existing in such data.

The second reason is that with the advance of computer technology and the resulting expansion of computer ability, improved visualization techniques of data, results of data analysis, and the features of data analysis have flourished. The features of complex data analysis involving imprecision and uncertainty have witnessed a crystallization of the exploratory visualization techniques for data.

The third reason is the limited precision in the data. In order to obtain precision from real data, we need to make many assumptions about the latent data structures. Under the many assumptions necessary for representing real data structures, even if we obtain a precise result for the data, since no one can know the real data structure, we cannot prove that the assumptions actually represent the real data structure. Statistical data analysis assumes “systematic" uncertainty for observational data. The amount of systemization is represented by statistical distribution. The concept of exploratory data analysis (Turkey (1977)) where the emphasis is on the idea that real data structure is on the multiple aspects of the data and that a model (or a structure) is not assumed has been proposed. For exploratory data analysis, the concept of statistical science was an ideal solution. However, the essence of observed data is not always based on “systematic" uncertainty. The need for analysis allowing for the most comprehensive uncertainty has been increased. As an ideal solution for analysis capturing unique features of data with the intention of discovering uncertainty and a set of robust and modern methods, fuzzy data analysis has been proposed.

From this background, the role of fuzzy data analysis techniques is growing and their potential is fully accepted in real world data analysis. In fuzzy data analysis, non-linear generalized models, symbolic data analysis, and kernel method are new and powerful methods that exhibit further progress with substantial reliance on the traditional statistical data analysis. In this chapter, we describe two kinds of innovative fuzzy data analysis techniques which exploit these methods hybrid with ordinary fuzzy clustering and principal component analysis.

The first presents a family of fuzzy clustering models based on a new aggregation operator defined on a product space of linear spaces. The purpose of this new aggregation operator is to control the variability of the similarity of objects in the fuzzy clustering model. In order to consider this variability, we have proposed an exploit of an aggregation operator. Although this aggregation operator can represent the variety of the common degree of belongingness of a pair of objects to pairs of fuzzy clusters, the representation of the variability of similarity still has a constraint, that is, the similarity still has to be explained by a linear structure with respect to pairs of fuzzy clusters. This is caused by the metric constraint and the fact that the aggregation operator is a binary operator. In order to solve this problem, we require a new definition of a function family on a product space of linear spaces and have similar conditions of the aggregation operator to represent the variability of similarity. Therefore, we propose a new aggregation operator called a generalized aggregation operator and present the better performance of the results of fuzzy clustering models based on the proposed aggregation operator.

The second proposal is a new correlation based on a fuzzy clustering result and a new principal component analysis (PCA) based on eigenvalues of this correlation. The merit of the use of fuzzy clustering based correlation to the PCA is that we can obtain the eigenvalues of the covariance matrix of variables for high-dimension and low-sample size (HDLSS) data in which the number of variables (dimensions) is much larger than the number of objects. Therefore, ordinary PCA can apply to HDLSS data by using fuzzy clustering based correlation.

In addition, we show a numerical example using micro array data, which is typical HDLSS data, to show a better performance of the proposed PCA with the fuzzy clustering based correlation when compared with ordinary PCA. Related with the PCA using the fuzzy clustering based correlation, there is a different PCA to apply to HDLSS data. While, many clustering techniques for interval-valued data have been proposed, especially for high dimension low sample-size interval-valued data, there has been no proposal for a variable selection added fuzzy clustering method. This chapter proposes this novel fuzzy clustering method for interval-valued data with an adaptable variable selection and proposes PCA based on this fuzzy clustering method.

There are three reasons why the method is necessary: First, our target data in this study is high dimension low sample-size data. Due to the curse of dimensionality, we tend to obtain a poor classification result for this type of data. The main cause of this is noise occurring from irrelevant and redundant variables (dimensions). Therefore, we need to use an adaptable variable selection to reduce or summarize variables. Second, the merit of fuzzy clustering is to obtain the results with uncertain cluster boundaries, which is well adjusted with the uncertainty situation of classification to data. This gives a more robust result for the noise in the data when compared with hard clustering while mathematically we can obtain a result with continuous values. Third, an adaptable representation of interval-valued data can be exploited to transform the original data into more manageable data in order to avoid the curse of dimensionality. Numerical examples show a high performance for the proposed method.

2. Fuzzy Clustering Model

2.1. Introduction

We have developed a nonlinear fuzzy clustering model (Sato-Ilic (2010a)) in order to deal with noisy data which is an extended model of a kernel fuzzy clustering model (Sato-Ilic et al. (2009)), an additive fuzzy clustering model (Sato et al. (1997)), and an additive clustering model (Shepard and Arabie (1979)). This model is one example of model-based clustering (clustering model) which is a category of clustering techniques whose essential feature is the assumption of a structure in the data. Through this feature, the mathematical properties of the obtained result tend to be clearer when compared with non-model-based clustering. In this model, we assume that all objects have some common properties and each common property is defined as a fuzzy cluster. That is, the similarity between a pair of objects is assumed to consist of some shared common properties of the objects. The shared common property is defined as common degree of belongingness of a pair of objects to a fuzzy cluster. Therefore, we do not need to define the metric to represent the similarity of objects. Since the difference of the definition of the metric causes the different clustering results, avoiding any definitions of metric has a benefit for the clustering. Exploiting this property has been discussed in certain areas such as genetics which have some merit for utilizing this model (ter Braak et al. (2010)). Also many algorithms related with this model have been developed through out several areas (Pedrycz et al. (2004), ter Braak et al. (2009), Runkler and Steinke (2010)).

Since an additive clustering model and an additive fuzzy clustering model have been discussed in a framework of “additive" clustering models, the common degree of belongingness of a pair of objects to a fuzzy cluster independently contributes to the similarity between the objects, so the interaction of a pair of objects for “different" fuzzy clusters cannot be considered. Therefore, we have taken the perspective of a nonlinear relationship among fuzzy clusters and have proposed a kernel fuzzy clustering model by extending the additive fuzzy clustering model for explaining the complexity of the noisy data.

However, in the kernel fuzzy clustering model, although the variety of the common degree of belongingness of a pair of objects to a fuzzy cluster is adjusted by using an aggregation operator, the variety of the obtained similarity cannot be explained due to the fact that the aggregation operator is a binary operator. Since the obtained similarity has various structures, in order to propose a general-purpose clustering model, the fuzzy clustering model based on operators on a product space of linear spaces which is inclusive of the adaptable variety is deemed indispensable.

Therefore, with the goal of creating a general-purpose clustering model having the merit of the degree of sharing common properties of objects, we propose a universal fuzzy clustering model with an implicit, internalized variability of the similarity structure. First, in order to implement the variability of the obtained similarity structure, we introduce a generalized aggregation operator (Sato-Ilic (2010a)) which is defined as a function on a product space of linear spaces and have similar conditions as aggregation operators.

Such a function has been discussed mathematically as a metric on the product space considering a probabilistic space (Tardiff (1976, 1980)). In addition, several definitions of multidimensional aggregation operators (Beliakov (2003), Ai-Ping et al. (2007), Merigo and Casanovas (2010)) have been proposed. Since we exploit the merit of the fuzzy clustering model in which we do not need to use any metric and our proposed generalized aggregation operator satisfies conditions similar to those of the aggregation operator which has suitable conditions for the clustering model in which it takes advantage of the property of degree of belongingness, we use the generalized aggregation operator in the proposed model.

This section consists of seven subsections. The following subsection describes the additive clustering model. Then in Section 2.3, we state an additive fuzzy clustering model. In Section 2.4, a kernel fuzzy clustering model is described. Section 2.5 proposes a fuzzy clustering model based on operators on a product space of linear spaces. Section 2.6 shows several numerical examples and Section 2.7 contains the concluding comments.

2. Additive Clustering Model

The additive clustering model (Shepard and Arabie (1979)) is defined as follows:

$$s_{ij} = \sum_{k=1}^{K} w_k p_{ik} p_{jk} + \varepsilon_{ij} \ , \qquad (1)$$

where $s_{ij}\ (i, j = 1, 2, \ldots, n)$ is a similarity data between objects i and j, K is the number of clusters, and w_k is a weight representing the salience of the property corresponding to the cluster k. n is the number of objects and ε_{ij} is an error. p_{ik} shows the status of belongingness of an object i to a cluster k. If an object i has the property of a cluster k, then $p_{ik} = 1$, otherwise it is 0. Therefore, p_{ik} satisfies the following condition:

$$p_{ik} \in \{0,1\}, \quad \forall i, k, \tag{2}$$

and the product $p_{ik} p_{jk}$ is unity only if both objects i and j belong to the cluster k. From the condition (2), it would be allowed that an object belongs to multiple clusters simultaneously. The cluster is defined, in this model, as a subset of all objects in which the objects included in the cluster share a common property. When a pair of objects has some common properties, this model assumes that these common properties “additively" contribute to the similarity between the pair of objects. That is, the degree of contribution of each common property to the similarity is mutually independent. For example, if a pair of objects i and j together belong to clusters $l_1, l_2, \ldots, l_m$, then the similarity s_{ij} is represented by the sum of weights of these clusters as follows:

$$s_{ij} = w_{l_1} + w_{l_2} + \cdots + w_{lm} + \varepsilon_{ij}.$$

Therefore, the similarity is represented by the degree of shared common properties.

2.3. Additive Fuzzy Clustering Model

From Eq. (3), it can be seen that since the similarity s_{ij} is observed as continuous values, in order to obtain the better fitness in which ε_{ij} is substantially small, the number of clusters tends to increase to explain observed similarity. In order to solve this problem, the additive fuzzy clustering model has been proposed. The additive fuzzy clustering model (Sato et al. (1997)) is defined as follows:

$$s_{ij} = \varphi\left(\rho_{ij}\right) + \varepsilon_{ij}, \tag{4}$$

where,

$$\rho_{ij} = \left(\rho\left(u_{i1}, u_{j1}\right), \ldots, \rho\left(u_{iK}, u_{jK}\right)\right) \in R^K. \tag{5}$$

Suppose that there exist K fuzzy clusters on a set of n objects, that is, the partition matrix $\mathbf{U} = \left(u_{ik}\right)$ is assumed to exist under the following conditions:

$$\sum_{k=1}^{K} u_{ik} = 1, \quad i = 1, \ldots, n, \tag{6}$$

$$u_{ik} \in [0,1], \quad i=1,\ldots,n, \quad k=1,\ldots,K \tag{7}$$

where u_{ik} shows a degree of belongingness of an object i to a cluster k. The purpose of model (4) is to estimate u_{ik} which minimize the sum of squared errors ε_{ij}. Let $\rho(u_{ik}, u_{jk})$ be a common degree of belongingness of a pair of objects i and j to a cluster k, namely, a degree of shared common property. To state simply, we assume that if all of $\rho(u_{ik}, u_{jk})$ are multiplied by α, then the similarity is also multiplied by α. Therefore, the function φ itself must satisfy the condition "positively homogeneous of degree 1 in the ρ", that is,

$$\alpha\varphi(\rho_{ij}) = \varphi(\alpha\rho_{ij}), \quad \alpha > 0. \tag{8}$$

We consider the following function as a typical function of φ:

$$s_{ij} = \varphi(\rho_{ij}) + \varepsilon_{ij} = \left\{\sum_{k=1}^{K} \rho^r (u_{ik}, u_{jk})\right\}^{\frac{1}{r}} + \varepsilon_{ij}, \quad 0 < r < +\infty. \tag{9}$$

We will deal with (9) ($r = 1$) hereafter, that is,

$$s_{ij} = \sum_{k=1}^{K} \rho(u_{ik}, u_{jk}) + \varepsilon_{ij} \tag{10}$$

The degree ρ is the aggregation operator satisfied the following conditions defined in Definition 1.

Definition 1: An aggregation operator (AO) is a binary operator ρ on the unit interval [0,1], that is a function ρ : $[0,1]\times[0,1] \to [0,1]$, such that $\forall a,b,c,d \in [0,1]$ $a,b,c,d \in R$, the following conditions are satisfied:

$$\rho(a,0) = \rho(0,a) = 0, \quad \rho(a,1) = \rho(1,a) = a,$$

$$\rho(a,c) \le \rho(b,d), \quad \text{whenever } a \le b,\ c \le d,$$

$$\rho(a,b) = \rho(b,a),$$

where $[0,1]\times[0,1]$ shows a product space. The first condition denotes the boundary condition which means that if one object belongs to a cluster completely, then the common degree of belongingness to the cluster equals the degree of the other object to the cluster, and if one object does not belong to the cluster, then it is 0. The second condition shows the condition of monotonicity, that is, the greater the degree of belongingness of objects to a cluster then the greater the common degree of belongingness of the objects. The third condition means the condition of symmetry which is that the common degree of belongingness of objects i and j is equivalent to the common degree of objects j and i. T-norm (Menger (1942), Schweizer and Sklar (1983)) is a typical example which satisfies the conditions in Definition 1. From Eqs. (6) and (7), we assume that $s_{ij} \in [0,1]$ in model (10). Algebraic product is an example of the t-norm, so if we assume as

$$\rho(u_{ik}, u_{jk}) = u_{ik}u_{jk}, \tag{11}$$

then the model (10) is represented as follows:

$$s_{ij} = \sum_{k=1}^{K} u_{ik}u_{jk} + \varepsilon_{ij} \tag{12}$$

In this model, if we put

$$u_{ik} = \sqrt{w_k}\, p_{ik}, \tag{13}$$

then the additive fuzzy clustering model shown in Eq. (12) is reduced to be the additive clustering model shown in Eq. (1). Therefore, the additive clustering model is a special case of the additive fuzzy clustering model which in turn is an extended model of the additive clustering model. Moreover, if we assume Eq. (13) which shows the additive clustering model, from Eq. (2), u_{ik} will have only two values for all i as follows:

$$u_{ik} \in \left\{0, \sqrt{w_k}\right\}, \quad \forall i. \tag{14}$$

This means that the flexibility of the representation of the fuzzy clustering result shown in Eq. (7) is substantially reduced to the two values shown in Eq. (14) when we use the additive clustering model. Therefore, the additive fuzzy clustering model can obtain a more flexible result by using fewer numbers of clusters when compared with the additive clustering model. This is caused by the change of the condition from Eq. (2) to Eq. (7), which shows the change from the hard clustering model to the fuzzy clustering model. Therefore, by introducing the concept of fuzzy logic to the additive clustering model, we can obtain a more flexible result. A practical algorithm for obtaining the solutions of the fuzzy clustering models is described in the literature. (Sato et al. (1997), Sato-Ilic and Jain (2006), ter Braak (2009))

2.4. Nonlinear Fuzzy Clustering Model

Since the additive fuzzy clustering model shown in Eq. (10) assumes the mutual independence among shared common properties, the interaction of different fuzzy clusters, which is a degree of shared common properties of objects to “different" fuzzy clusters, cannot be reflected in this model. Therefore, given the noisy data and the models lack of power to produce an explanation, the result tends to be imprecise. In order to overcome this problem, one simple idea is to consider $\rho(u_{ik}, u_{jl}), k \neq l$ as a common degree of belongingness of a pair of objects i and j to clusters k and l and put this to

-

-

-

Bibliography

Ahn, J., Marron, J.S., Muller, K.M., Chi, Y-Y. (2007) "The High-Dimension, Low-Sample-Size Geometric Representation Holds under Mild Conditions," *Biometrika*, 94, 3, 760-766. [This paper discusses geometric representation of high-dimension and low-sample size data]

Ai-Ping, L., Yan, J., Quan-Yuan, W. (2007) "Harmonic Triangular Norm Aggregation Operators in Multicriteria Decision Systems," *Journal of Convergence Information Technology*, 2, 1, 83-92. [This paper proposes aggregation operators for multicriteria decision]

Baik, J., Arous, G.B., Peche, S. (2005) "Transition of the Largest Eigenvalue for Nonnull Complex Sample Covariance Matrices," *The Annals of Probability*, 33, 5, 1643-1697.[This paper discusses of features of largest eigenvalue of sample covariance matrix]

Beliakov, G, (2003) "How to Build Aggregation Operators from Data," *International Journal of Intelligent Systems*, 18, 903-923. [This paper discusses aggregation operators for the multiple elements of vectors]

Bezdek, J.C. (1981) *Pattern Recognition with Fuzzy Objective Function Algorithms*, Plenum, New York. [This book discusses fuzzy c-means method]

Billard, L., Diday, E. (2007) *Symbolic Data Analysis: Conceptual Statistics and Data Mining*, Wiley. [This book discusses symbolic data analysis]

Bock, H.H., Diday E. (Eds.) (2000) *Analysis of Symbolic Data*, Springer. [This book discusses symbolic data analysis]

Friedman, J.H., Meulman, J.J. (2004) "Clustering Objects on Subsets of Attributes," *Journal of the Royal Statistical Society*, Series B, 66, 4, 815-849. [This paper proposes the COSA method]

Gower, J.C. (1966) "Some Distance Properties of Latent Roots and Vector Methods used in Multivariate Analysis," *Biometrika*, 53, 325-338. [This paper discusses multidimensional scaling]

Hall, P., Marron, J.S., Neeman, A. (2005) "Geometric Representation of High Dimension Low Sample Size Data," *Journal of Royal Statistical Society*, 67, Part 3, 427-444. [This paper discusses distance of a pair of objects for high-dimension and low-sample size data]

Hastie, T., Tibshirani, R., Friedman, J. (2009) *The Elements of Statistical Learning*, 2nd ed., Springer. [This book discusses statistical learning techniques]

Johnstone, I.M. (2001) "On the Distribution of the Largest Eigenvalue in Principal Components Analysis," *The Annals of Statistics*, 29, 2, 295-327. [This paper discusses largest eigenvalue in PCA]

Jolliffe, I.T. (2002) *Principal Component Analysis*, 2nd ed., Springer. [This book explains PCA]

Kaufman, L., Rousseeuw, P.J. (1990) *Finding Groups in Data*, John Wiley & Sons. [This book explains the FANNY method in fuzzy clustering]

Kruskal, J.B., Wish, M. (1978) *Multidimensional Scaling*, Sage Publications. [This book explains multidimensional scaling]

Menger, K. (1942) "Statistical Metrics," *Proc. Nat. Acad. Sci.*, USA, 28, 535-537. [This paper proposes T-norm]

Merigo, J.M., Casanovas, M. (2010) "Fuzzy Generalized Hybrid Aggregation Operators and its Application in Fuzzy Decision Making," *International Journal of Fuzzy Systems*, 12, 1, 15-24. [This paper discusses aggregation operator in fuzzy decision]

Pedrycz, W., Loiac, V., Senatore, S. (2004) "Proximity based Fuzzy Clustering (P-FCM),"*Fuzzy Sets and Systems*, 148, 21-41. [This paper presents the p-fcm method]

Runkler, T.A., Steinke, F. (2010) "A New Approach to Clustering Using Eigen Decomposition," *WCCI 2010 IEEE World Congress on Computational Intelligence*, 306-311. [This paper presents fuzzy clustering using eigen decomposition]

Sato, M., Sato, Y., Jain, L.C. (1997) *Fuzzy Clustering Models and Applications*, Springer. [This book explains fuzzy clustering models]

Sato-Ilic, M. (2004) "Self-Organized Fuzzy Clustering," *Intelligent Engineering Systems through Artificial Neural Networks*, 14, 579-584. [This paper proposes self-organized fuzzy clustering method]

Sato-Ilic, M., Kuwata, T. (2005) "On Fuzzy Clustering based Self-Organized Methods," *FUZZ-IEEE2005*, 973-978. [This paper presents fuzzy clustering based on self-organized methods]

Sato-Ilic, M., Jain, L.C. (2006) *Innovations in Fuzzy Clustering*, Springer. [This book explains several new fuzzy clustering methods]

Sato-Ilic, M. (2008) "Fuzzy Variable Selection with Degree of Classification based on Dissimilarity between Distributions of Variables," *International Journal of Intelligent Technology and Applied Statistics*, 1, 2, 1-18. [This paper proposes a fuzzy variable selection method]

Sato-Ilic, M., Ito, S., Takahashi, S. (2009) Nonlinear Kernel-Based Fuzzy Clustering Model, *Developments in Fuzzy Clustering*, D.A. Viattchenin ed., VEVER, Minsk (Belarus), 56-73. [This paper presents the nonlinear kernel-based fuzzy clustering model]

Sato-Ilic, M. (2010a) "Generalized Aggregation Operator Based Nonlinear Fuzzy Clustering Model," *Intelligent Engineering Systems through Artificial Neural Networks*, 20, 493-500. [This paper proposes generalized aggregation operator based nonlinear fuzzy clustering model]

Sato-Ilic, M. (2010b) "A Cluster-Target Similarity Based Principal Component Analysis for Interval-Valued Data," *th International Conference on Computational Statistics*, Physica-Verlag, 1605-1612. [This paper proposes clustering-based PCA for interval-valued data]

Sato-Ilic, M. (2011a) "Symbolic Clustering with Interval-Valued Data," *Procedia Computer Sciences, Elsevier*, 6, 358-363. [This paper proposes a symbolic clustering method for interval-valued data]

Sato-Ilic, M. (2011b) "Clustering High Dimension Low Sample-Size Data with Fuzzy Cluster-based Principal Component Analysis," *the 58th ISI World Statistics Congress*, cps024-5. [This paper discusses fuzzy clustering-based PCA for HDLSS data]

Sato-Ilic, M. (2012a) "Structural Classification based Correlation and its Application to Principal Component Analysis for High-Dimension Low-Sample Size Data," *IEEE World Congress on Computational Intelligence*, 981-988. [This paper proposes a new cluster-based correlation and its use for the HDLSS data]

Sato-Ilic, M. (2012b) "On Fuzzy Clustering Based Correlation," *Procedia Computer Sciences, Elsevier*, 12, 230-235. [This paper proposes fuzzy clustering based correlation]

Schweizer, B., Sklar, A. (1983) *Probabilistic Metric Spaces*, North-Holland. [This book discusses T-norms]

Shawe-Taylor, J., Cristianini, N. (2004) *Kernel Methods for Pattern Analysis*, Cambridge University Press. [This book explains kernel methods]

Shepard, R.N., Arabie, P. (1979) "Additive Clustering: Representation of Similarities as Combinations of Discrete Overlapping Properties," *Psychological Review*, 86, 2, 87-123. [This paper proposes the additive clustering model]

Tardiff, R.M. (1976) "Topologies for Probabilistic Metric Spaces," *Pacific Journal of Mathematics*, 65, 1, 233-251. [This paper discusses topologies for probabilistic metric spaces]

Tardiff, R.M. (1980) "On a Functional Inequality Arising in the Construction of the Product of Several Metric Spaces," *Aequationes Mathematicae*, 20, 51-58. [This paper discusses of product of metric spaces]

ter Braak, C.F.J., Kourmpetis, Y., Kiers, H.A.L., Bink, M.C.A.M. (2009) "Approximating a Similarity Matrix by a Latent Class Model: A Reappraisal of Additive Fuzzy Clustering," *Computational Statistics and Data Analysis*, 53, 3813-3193. [This paper discusses algorithm for additive fuzzy clustering model]

ter Braak, C.F.J., Boer, M.P., Totir, L.R., Winkler, C.R., Smith, O.S., Bink, M.C.A.M. (2010) "Identity-by-Descent Matrix Decomposition Using Latent Ancestral Allele Models," *Genetics*. [This paper discusses application of additive fuzzy clustering model]

Tukey, J.W. (1977) *Exploratory Data Analysis*, Addison-Wesley Publishing. [This book explains exploratory data analysis]

Welsh, J.B., Sapinoso, L.M., Su, A.I., Kern, S.G., Wang-Rodriguez, J., Moskaluk, C.A., Frierson, Jr., H.F., Hampton, G.M. (2001) "Analysis of Gene Expression Identifies Candidate Markers and Pharmacological Targets in Prostate Cancer," *Cancer Research*, 61, 5974-5978. [This paper discusses methods for gene expression identifiers]

Biographical Sketches

Prof. Mika Sato-Ilic received a Doctorate of Engineering (PhD) from Hokkaido University, specializing in Intelligent Data Analysis. She currently holds the position of Professor in the Faculty of Engineering, Information and Systems, at the University of Tsukuba, Japan. She is the founding editor-in-chief of the International Journal of Knowledge Engineering and Soft Data Paradigms, Associate Editor of Neuro computing, Regional Editor of International Journal on Intelligent Decision Technologies and Associate Editor of the International Journal of Innovative Computing, Information and Control Express Letters, as well as serving on the editorial board of several other journals. In addition, she is currently a council member of the International Association for Statistical Computing which is a section of the International

Statistical Institute (ISI) and a Senior Member of the IEEE where she holds several positions including the Vice-Chair of the Fuzzy Systems Technical Committee of the IEEE Computational Intelligence Society. In addition, she has served on several IEEE committees including the administration committee, vice program chair, special sessions co-chair. Her academic output includes 4 books, 7 book chapters and over 100 journal and conference papers. Her research interests include the development of methods for data mining, multi-dimensional data analysis, multi-mode multi-way data theory, pattern classification, and computational intelligence techniques for which she has received several academic awards.