A Cloud Business Intelligence System for Visual Analytics with Big Data

Chia-Hui Huang, Keng-Chieh Yang, and Han-Ying Kao

***Abstract*—Big data has become one of new research frontiers. It is a collection of a large-scale and complex data sets that it becomes more difficult to process using current database management systems and traditional data processing applications. There are two challenges while dealing with big data: (1) how to analyze big data efficiently; (2) visualization and presentation of big data because of the larger volume, variety, and velocity of the information.**

This study proposes a cloud business intelligence system for visual analytics with big data. A new kernel method for analyzing big data is proposed. The principle of semismooth support vector machine is introduced to collaborate with the interval regression model. The proposed kernel method can resolve the following problem efficiently: (1) big data; (2) noises and interaction of the separation margin; (3) unbalance of the separation margin.

***Index Terms*—big data, data visualization, cloud business intelligence system, interval regression model.**

I. Introduction

INFORMATION or data management according to a series of studies by Carnegie Mellow University entails the process of collecting, classifying, storing, retrieving, and managing data.Data collected from different processes is used to make decisions feasible to the understanding and requirements of those executing and consuming the result of the process. Data warehousing evolved to support the decision-making process of being able to collect, store, and manage data, applying statistical methods of measurement to create a report and analysis platform.

Big data has become one of new research frontiers. Generally speaking, big data is a collection of a large-scale and complex data sets that it becomes more difficult to process using current database management systems and traditional data processing applications. In 2012, Gartner Inc. gave a definition of big data as: "*Big data is high volume, high velocity, and/or high variety information assets that require new forms of processing to enable enhanced decision making, insight discovery and process optimization*" [22]. The trend of big data sets is due to the additional information derivable from analysis of a large set of related data, as compared to separate smaller sets with the same total amount of data.

There are two challenges while dealing with big data:

(1) Analyzing big data: One of the major applications of the future parallel, distributed, and cloud systems is in big data analytic [1], [5], [12], [13], [18], [21], [23], [27]. Most concerned issues are dealing with big data sets which often require computation resources provided by public cloud services. How to analyze big data efficiently becomes a big challenge [14], [18], [35], [38], [40].

(2) Visualization and presentation of big data: The other challenge is the visualization and presentation of big data because of the larger volume, variety, and velocity of the information [2], [6], [11], [19], [20]. Presentation in visual form provides the capability to summarize various amounts of information into a format that may provide the ability to drill down to further level of detail. Being able to drill down from higher level data to more granular detail is a very important feature in dashboard design. With the feature of drill down analysis, users can easily navigate through the different levels of data.

Although a large part of data produced nowadays is unstructured, relational databases have been the choice most organisations have made to store data about their customers, sales, and products. As data managed by traditional database management ages, it is moved to data warehouses for analysis and for sporadic retrieval. Models such as MapReduce are generally not the most appropriate to analyse such relational data. Attempts have been made to provide hybrid solutions that incorporate MapReduce to perform some of the queries and data processing required by database management [1].

There are many research challenges in the field of big data visualisation [2], [6], [11], [19], [20], [29], [32]. More efficient data processing techniques are required in order to enable real-time visualisation. Andrienko et al. [2] proposed interactive visual display for analysis of movement behaviour of people, vehicle, and animals. The visualisation tool displays the movement data, information about the time spent in a place, and the time interval from one place to another. Choo and Park [6] appointed some techniques that can be employed with this objective, such as reduction of accuracy of results, coarsely processing of data points, compatible with the resolution of the visualisation device, reduced convergence, and data scale confinement. Methods considering each of these techniques could be further researched and improved.

This study proposes a cloud business intelligence system for visual analytics with big data. A new kernel method for analyzing big data is proposed. The principle of semismooth support vector machine is introduced to collaborate with the interval regression model. The rest of the study is organized as follows. Section II proposes the kernel for analyzing big data. Section III presents our discussions and expected outcomes.

Manuscript received December 29, 2018; revised January 28, 2019. This work was supported in part by the Ministry of Science and Technology, Taiwan, for financially supporting this research under grant No. MOST 104-2410-H-141-016-.

Chia-Hui Huang is with the Department of Business Administration, National Taipei University of Business, Taiwan. e-mail:leohuang@ntub.edu.tw.

Keng-Chieh Yang is with the Department of Finance and Information, National Kaohsiung University of Science and Technology, Taiwan. Han-Ying Kao is with the Department of Computer Science and Information Engineering, National Dong Hwa University, Taiwan.

II. Proposed Methods

One of the major applications of the future parallel, distributed, and cloud systems is in big data analytic [5], [18], [23], [35]. Most concerned issues are dealing with large-scale sets which often require computation resources provided by public cloud services. How to analyze big data efficiently becomes a big challenge.

The support vector machine (SVM) has shown to be an efficient approach for a variety of data mining, classification, analysis, pattern recognition, and distribution estimation [3], [24], [26], [28], [33], [34], [37], [39]. Recently, using SVM to solve the interval regression model [36] has become an alternative approach. Hong and Hwang [15] evaluated interval regression models with quadratic loss SVM. Bisserier et al. [4] proposed a revisited fuzzy regression method where a linear model is identified from Crisp-Inputs Fuzzy-Outputs (CISO) data. D'Urso et al. [8] presented fuzzy clusterwise regression analysis with LR fuzzy response variable and numeric explanatory variables. The suggested model is to allow for linear and non-linear relationship between the output and input variables. Huang [16] solved interval regression model with reduced support vector machine. Jeng et al. [17] developed a support vector interval regression networks (SVIRNs) based on both SVM and neural networks.

However, there are several main problems while using SVM model:

(1) Big data: when dealing with big data sets, the solution by using SVM with a nonlinear kernel may be difficult to find.
(2) Noises and Interaction: the distribution of data becomes hard to be described and the separation margin between classes becomes a "gray" zone.
(3) Unbalance: the number of samples from one class is much larger than the number of samples from other classes. It causes the excursion of separation margin.

Under this circumstance, developing an efficient method to analyze big data becomes important. The semismooth support vector machine has been proved more efficient than the traditional SVM in processing large-scale data [9]. The main idea of semismooth support vector machine is implemented the class of a semismooth function [7], [30], [31] and has been extended to nonlinear separation surfaces by using a nonlinear kernel technology. A semismooth method based on the Fischer-Burmeister merit function [10] is formulated to solve the linear mixed complementarity problems defined by the necessary and sufficient first-order optimality conditions. The semismooth method reformulates the complementarity problem as a system of nonlinear, nonsmooth equations, and applies a generalized Newton method to find a solution.

Suppose that m training data $\{x_i, y_i\}$, $i = 1, 2, \ldots, m$ are given, where $x_i \in \Re^n$ are the input patterns and $y_i \in \{-1, 1\}$ are the related target values of two-class pattern classification case. Then the standard support vector machine with a linear kernel [39] is

$$\begin{aligned} &\min_{w,b,\xi} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m}\xi_i^2 \\ &s.t. \\ &\quad y_i(w^t x_i + b) \geq 1 - \xi_i \\ &\quad \xi_i \geq 0, i = 1, 2, \ldots, m \end{aligned} \tag{1}$$

where b is the location of hyperplane relative to the origin. The regularization constant C is a positive parameter to control the tradeoff between the training error and the part of maximizing the margin that is achieved by minimizing $\|w\|^2$. ξ_i is the slack variable with weight $C/2$. $\|w\|$ is the Euclidean norm of w which is the normal to the following hyperplanes

$$w^t x_i + b = +1, \text{ for } y_i = +1 \tag{2}$$

$$w^t x_i + b = -1, \text{ for } y_i = -1 \tag{3}$$

The first hyperplane (2) bounds the class $\{+1\}$ and the second hyperplane (3) bounds the class $\{-1\}$. The linear separating hyperplane is

$$w^t x_i + b = 0 \tag{4}$$

A semismooth function is a generalized notion of continuously differentiable functions that are both Lipschitzian and directionally differentiable. Let $G : \Re^n \mapsto \Re^n$ be a Lipschitzian function and D_G denote the set of points where G is differentiable. In Ferris and Munson's approach [9], b^2 is added to the objective function of (1). This is equivalent to adding a constant feature to the training data and finding a separating hyperplane through the origin.

$$\begin{aligned} &\min_{w,b,\xi} \frac{1}{2}(\|w\|^2 + 2b^2) + C\sum_{i=1}^{m}\xi_i^2 \\ &s.t. \\ &\quad y_i(w^t x_i + b) \geq 1 - \xi_i \\ &\quad \xi_i \geq 0, i = 1, 2, \ldots, m \end{aligned} \tag{5}$$

where $\xi_i = \{1 - y_i(w^t x_i + b)\}_+$ for all i and the "+" function is defined as $x_+ = \max\{0, x\}$. Then (1) can be reformulated as the following minimization problem by replacing ξ_i with $\{1 - y_i(w^t x_i + b)\}_+$

$$\min_{w,b} \frac{1}{2}(\|w\|^2 + 2b^2) + C\sum_{i=1}^{m}\{1 - y_i(w^t x_i + b)\}_+^2 \tag{6}$$

The "+" function in semismooth support vector machine is approximated by a semismooth function, $G(x, \rho)$, as follows

$$G(x, \rho) = x + \rho \log(1 + e^{-\rho x}), \rho > 0 \tag{7}$$

where $\rho > 0$ is the semismooth parameter. The $G(x, \rho)$ with a smoothing parameter ρ is to replace the "+" function of (6) to obtain the following semismooth support vector machine with a linear kernel

$$\min_{w,b} \frac{1}{2}(\|w\|^2 + 2b^2) + C\sum_{i=1}^{m} G(\{1 - y_i(w^t x_i + b)\}, \rho)^2 \tag{8}$$

For specific data sets, an appropriate nonlinear mapping $x \mapsto \phi(x)$ can be used to embed the original $\Re^n$ features into a Hilbert feature space $\mathcal{F}$, $\phi : \Re^n \mapsto \mathcal{F}$, with a nonlinear kernel $K(x_i, x_j) \equiv \phi(x_i)^t \phi(x_j)$. Thus (8) can be extended to the semismooth support vector machine with a nonlinear kernel

$$\begin{aligned} &\min_{w,b} \frac{1}{2}(\|w\|^2 + 2b^2) + \\ &\quad C\sum_{i=1}^{m} G(\{1 - y_i(\sum_{j=1}^{m} v_j K(x_i, x_j) + b)\}, \rho)^2 \end{aligned} \tag{9}$$

where $\sum_{j=1}^{m} v_j K(x_i, x_j) + b$ is the nonlinear semismooth support vector machine classifier. The coefficient v_j is determined by solving an optimization problem (9) and the data points with corresponding non-zero coefficients.

With the principle of semismooth support vector machine, we can formulate the interval linear regression model as follows

$$\begin{aligned}
&\min_{\bar{a},\bar{c},\bar{d}} \frac{1}{2}(\bar{a}^t\bar{a} + \bar{c}^t\bar{c} + \bar{d}^t\bar{d} + 2b^2) + \\
&\quad C\sum_{i=1}^{m} G(\{1 - y_i(w^t x_i + b)\}, \rho)^2 \\
&s.t. \\
&\quad \bar{a}\boldsymbol{x}_i + \bar{c}|\boldsymbol{x}_i| + \bar{d}|\boldsymbol{x}_i| \geq y_i + e_i \\
&\quad \bar{a}\boldsymbol{x}_i - \bar{c}|\boldsymbol{x}_i| - \bar{d}|\boldsymbol{x}_i| \leq y_i - e_i \\
&\quad \bar{a}\boldsymbol{x}_i + \bar{c}|\boldsymbol{x}_i| \leq y_i + e_i \\
&\quad \bar{a}\boldsymbol{x}_i - \bar{c}|\boldsymbol{x}_i| \geq y_i - e_i \\
&\quad i = 1, 2, \ldots, m
\end{aligned} \tag{10}$$

where $\bar{a}$, $\bar{c}$, and $\bar{d}$ are the collections of all a_i, c_i, and d_i, $i = 1, 2, \ldots, m$, respectively.

Given (10), the corresponding Lagrangian objective function is

$$\begin{aligned}
L :=& \frac{1}{2}(\bar{a}^t\bar{a} + \bar{c}^t\bar{c} + \bar{d}^t\bar{d} + 2b^2) \\
&+ C\sum_{i=1}^{m} G(\{1 - y_i(w^t x_i + b)\}, \rho)^2 \\
&- \sum_{i=1}^{m} \lambda_{1i}(y_i + e_i - \bar{a}\boldsymbol{x}_i - \bar{c}|\boldsymbol{x}_i|) \\
&- \sum_{i=1}^{m} \lambda_{2i}(\bar{a}\boldsymbol{x}_i - \bar{c}|\boldsymbol{x}_i| - y_i + e_i) \\
&- \sum_{i=1}^{m} \lambda_{3i}(\bar{a}\boldsymbol{x}_i + \bar{c}|\boldsymbol{x}_i| + \bar{d}|\boldsymbol{x}_i| - y_i - e_i) \\
&- \sum_{i=1}^{m} \lambda_{4i}(y_i - e_i - \bar{a}\boldsymbol{x}_i + \bar{c}|\boldsymbol{x}_i| + \bar{d}|\boldsymbol{x}_i|)
\end{aligned} \tag{11}$$

where L is Lagrangian and λ_{1i}, λ_{2i}, λ_{3i}, and λ_{4i} are Lagrange multipliers. The idea to construct a Lagrange function from the objective function and the corresponding constraints is to introduce a dual set of variables. It can be shown that the Lagrangian function has a saddle point with respect to the primal and dual variables in the solution [25].

The Karush-Kuhn-Tucker (KKT) conditions that the partial derivatives of L with respect to the primal variables $(\bar{a}, \bar{c}, \bar{d})$ for optimality

$$\frac{\partial L}{\partial \bar{a}} = 0 \Rightarrow \bar{a} = \sum_{i=1}^{m}(\lambda_{3i} - \lambda_{4i})\boldsymbol{x}_i - \sum_{i=1}^{m}(\lambda_{1i} - \lambda_{2i})\boldsymbol{x}_i \tag{12}$$

$$\begin{aligned}
\frac{\partial L}{\partial \bar{c}} = 0 \Rightarrow \bar{c} =& \sum_{i=1}^{m}(\lambda_{3i} + \lambda_{4i})|\boldsymbol{x}_i| \\
&- \sum_{i=1}^{m}(\lambda_{1i} + \lambda_{2i})|\boldsymbol{x}_i|
\end{aligned} \tag{13}$$

$$\frac{\partial L}{\partial \bar{d}} = 0 \Rightarrow \bar{d} = \sum_{i=1}^{m}(\lambda_{3i} + \lambda_{4i})|\boldsymbol{x}_i| \tag{14}$$

Substituting (12)–(14) in (11) yields the following optimization problem

$$\begin{aligned}
\max \frac{1}{2}(&\sum_{i,j=1}^{m}(\lambda_{1i} - \lambda_{2i})(\lambda_{1j} - \lambda_{2j})\boldsymbol{x}_i^t\boldsymbol{x}_j \\
&- \sum_{i,j=1}^{m}(\lambda_{3i} - \lambda_{4i})(\lambda_{3j} - \lambda_{4j})\boldsymbol{x}_i^t\boldsymbol{x}_j \\
&+ \sum_{i,j=1}^{m}(\lambda_{1i} + \lambda_{2i})(\lambda_{1j} + \lambda_{2j})|\boldsymbol{x}_i|^t|\boldsymbol{x}_j| \\
&- \sum_{i,j=1}^{m}(\lambda_{3i} + \lambda_{4i})(\lambda_{3j} + \lambda_{4j})|\boldsymbol{x}_i|^t|\boldsymbol{x}_j| \\
&+ \sum_{i,j=1}^{m}(\lambda_{2i} - \lambda_{2j})(\lambda_{3i} - \lambda_{3j})\boldsymbol{x}_i^t\boldsymbol{x}_j \\
&- \sum_{i,j=1}^{m}(\lambda_{2i} + \lambda_{2j})(\lambda_{3i} + \lambda_{3j})|\boldsymbol{x}_i|^t|\boldsymbol{x}_j| \\
&- 2\sum_{i,j=1}^{m}((\lambda_{3i} + \lambda_{4i})(\lambda_{3j} + \lambda_{4j})|\boldsymbol{x}_i|^t|\boldsymbol{x}_j| + b^2)) \\
&+ C\sum_{i=1}^{m} G(\{1 - y_i(w^t x_i + b)\}, \rho)^2 \\
s.t.& \\
&\lambda_{1i}, \lambda_{2i}, \lambda_{3i}, \lambda_{4i} \geq 0
\end{aligned} \tag{15}$$

Similarly, we can obtain the interval nonlinear regression model by mapping $x \mapsto \phi(x)$ to embed the original $\Re^n$ features into a Hilbert feature space $\mathcal{F}$, $\phi : \Re^n \mapsto \mathcal{F}$, with a nonlinear kernel $K(x_i, x_j) \equiv \phi(x_i)^t\phi(x_j)$. The optimization problem as (16) can be deduced by replacing $\boldsymbol{x}_i^t\boldsymbol{x}_j$ and $|\boldsymbol{x}_i|^t|\boldsymbol{x}_j|$ in (15) with $K(\boldsymbol{x}_i, \boldsymbol{x}_j)$ and $K(|\boldsymbol{x}_i|, |\boldsymbol{x}_j|)$, respectively

$$\begin{aligned}
\max \frac{1}{2}(&\sum_{i,j=1}^{m}(\lambda_{1i} - \lambda_{2i})(\lambda_{1j} - \lambda_{2j})K(\boldsymbol{x}_i, \boldsymbol{x}_j) \\
&- \sum_{i,j=1}^{m}(\lambda_{3i} - \lambda_{4i})(\lambda_{3j} - \lambda_{4j})K(\boldsymbol{x}_i, \boldsymbol{x}_j) \\
&+ \sum_{i,j=1}^{m}(\lambda_{1i} + \lambda_{2i})(\lambda_{1j} + \lambda_{2j})K(|\boldsymbol{x}_i|, |\boldsymbol{x}_j|) \\
&- \sum_{i,j=1}^{m}(\lambda_{3i} + \lambda_{4i})(\lambda_{3j} + \lambda_{4j})K(|\boldsymbol{x}_i|, |\boldsymbol{x}_j|) \\
&+ \sum_{i,j=1}^{m}(\lambda_{2i} - \lambda_{2j})(\lambda_{3i} - \lambda_{3j})K(\boldsymbol{x}_i, \boldsymbol{x}_j) \\
&- \sum_{i,j=1}^{m}(\lambda_{2i} + \lambda_{2j})(\lambda_{3i} + \lambda_{3j})K(|\boldsymbol{x}_i|, |\boldsymbol{x}_j|) \\
&- 2\sum_{i,j=1}^{m}((\lambda_{3i} + \lambda_{4i})(\lambda_{3j} + \lambda_{4j})K(|\boldsymbol{x}_i|, |\boldsymbol{x}_j|) + b^2)) \\
&+ C\sum_{i=1}^{m} G(\{1 - y_i(w^t x_i + b)\}, \rho)^2 \\
s.t.& \\
&\lambda_{1i}, \lambda_{2i}, \lambda_{3i}, \lambda_{4i} \geq 0
\end{aligned} \tag{16}$$

The followings are well-known nonlinear kernels, where σ, γ, r, h, and θ are kernel parameters.

ISBN: 978-988-14048-5-5
ISSN: 2078-0958 (Print); ISSN: 2078-0966 (Online)

(1) Gaussian (radial basis) kernel: $e^{\frac{-\|x_i - x_j\|^2}{2\sigma^2}}$, $\sigma > 0$. [28]
(2) Hyperbolic tangent kernel: $\tanh(\gamma x_i x_j^t + \theta)$, $\gamma > 0$. [34]
(3) Polynomial kernel: $(\gamma x_i x_j^t + r)^h$, $h \in \mathbb{N}$, $\gamma > 0$ and $r \geq 0$. [39]

III. Conclusion

This study proposes a cloud business intelligence system for visual analytics with big data. Cloud based BI benefits users by providing applications scalability. Cloud computing technology makes it easier to deliver the BI software as a service for users who want to acquire cloud BI without managing the hardware and software.

In addition a kernel method for analyzing big data is proposed. The principle of semismooth support vector machine is introduced to collaborate with the interval regression model. The main idea of semismooth support vector machine is implemented the class of semismooth functions and has been extended to nonlinear separation surfaces by using a nonlinear kernel technology. The proposed kernel method can resolve the following problem efficiently: (1) big data; (2) noises and interaction of the separation margin; (3) unbalance of the separation margin.

For future extensions, this work may be implemented into complete decision support systems which include various models and reasoning strategies. Through the user interfaces, the decision-makers may modify the network structures, relevant parameters and distributions. There are some potential themes:

- Potential applications of intelligent decision support with big data, such as supply chain management, business strategic analysis, and biomedicine.
- Integrated analysis with fuzzy sets and rough sets in various graphical decision model.
- Hybrid decision analysis with fuzzy sets and rough sets in cloud environment.

Acknowledgment

The author would like to appreciate the anonymous referees for their useful comments and suggestions which helped to improve the quality and presentation of this manuscript.

References

[1] D.J. Abadi, "Data Management in the Cloud: Limitations and Opportunities," *IEEE Data Engineering Bulletin*, vol. 32, no. 1, pp. 3–12, 2009.
[2] G. Andrienko, N. Andrienko and S. Wrobel, "Visual Analytics Tools for Analysis of Movement Data," *ACM SIGKDD Explorations Newsletter*, vol. 9, no. 2, pp. 38–46, 2007.
[3] M. Arun Kumar, R. hemchandani, M. Gopal and S. Chandra, "Knowledge based Least Squares Twin Support Vector Machines," *Information Sciences*, vol. 180, pp. 4606–4618, 2010.
[4] A. Bisserier, R. Boukezzoula and S. Galichet, "A Revisited Approach to Linear Fuzzy Regression using Trapezoidal Fuzzy Intervals," *Information Sciences*, vol. 180, pp. 3653–3673, 2010.
[5] C.L.P. Chen and C.Y. Zhang, "Data-intensive Applications, Challenges, Techniques and Technologies: A Survey on Big Data," *Information Sciences*, vol. 275, pp. 314–347, 2014.
[6] J. Choo and H. Park, "Customizing Computational Methods for Visual Analytics with Big Data," *IEEE Computer Graphics and Applications*, vol. 33, no. 4, pp. 22–28, 2013.
[7] T. De Luca, F. Facchinei and C. Kanzow, "A Semismooth Equation Approach to the Solution of Nonlinear Complementarity Problems," *IEEE Mathematical Programming*, vol. 75, pp. 407–439, 1996.
[8] P. D'Urso, R. Massari and A. Santoro, "Robust Fuzzy Regression Analysis," *Information Sciences*, vol. 181, pp. 4154–4174, 2011.
[9] M.C. Ferris and T.S. Munson, "Semismooth Support Vector Machines," *Mathematical Programming*, vol. 101, no. 1, pp. 185-204, 2004.
[10] A. Fischer, "A Special Newton-type Optimization Method," *Optimization*, vol. 24, pp. 269–284, 1992.
[11] D. Fisher, R. DeLine, M. Czerwinski and S. Drucker, "Interactions with Big Data Analytics," *Interactions*, vol. 19, no. 3, pp. 50–59, 2012.
[12] B. Franks, *Taming the Big Data Tidal Wave: Finding Opportunities in Huge Data Streams with Advanced Analytics*, Wiley and SAS Business Series, Wiley, 2012.
[13] K. Goodhope, J. Koshy, J. Kreps, N. Narkhede, R. Park, J. Rao and V.Y. Ye, "Building LinkedIns Real-Time Activity Data Pipeline," *Bulletin of the Technical Committee on Data Engineering*, vol. 35, no. 2, pp. 33–45, 2012.
[14] A. Guazzelli, K. Stathatos and M. Zeller, "Efficient Deployment of Predictive Analytics Through Open Standards and Cloud Computing," *ACM SIGKDD Explorations Newsletter*, vol. 11, no. 1, pp. 32–38, 2009.
[15] D.H. Hong and C.H. Hwang, "Interval Regression Analysis using Quadratic Loss Support Vector Machine," *IEEE Transactions on Fuzzy Systems*, vol. 13, pp. 229–237, 2005.
[16] C.H. Huang, "A Reduced Support Vector Machine Approach for Interval Regression Analysis," *Information Sciences*, vol. 217, pp. 56–64, 2012.
[17] J.T. Jeng, C.C. Chuang and S.F. Su, "Support Vector Interval Regression Networks for Interval Regression Analysis," *Fuzzy Sets and Systems*, vol. 138, pp. 283–300, 2003.
[18] K. Kambatla, G. Kollias, V. Kumar and A.Grama, "Trends in Big Data Analytics," *Journal of Parallel and Distributed Computing*, vol. 74, pp. 2561–2573, 2014.
[19] D. Keim, "Information Visualization and Visual Data Mining," *IEEE Transactions on Visualization and Computer Graphics*, vol. 8, no. 1, pp. 1–8, 2002.
[20] R. Kosara and J. Mackinlay, "Storytelling: the Next Step for Visualization," *Computer*, vol. 46, no. 5, pp. 44–50, 2013.
[21] A. Kumar, F. Niu and C. Ré, "Hazy: Making it Easier to Build and Maintain Big-Data Analytics," *Communications of the ACM*, vol. 56, no. 3, pp. 40–49, 2013.
[22] D. Laney, *The Importance of Big Data: A Definition*, Gartner Inc., 2012
[23] V. López, S. del Río, J.M. Benítez and F. Herrera, "Cost-Sensitive Linguistic Fuzzy Rule Based Classification Systems under The Mapreduce Framework for Imbalanced Big Data," *Fuzzy Sets and Systems*, vol. 258, pp. 5–38, 2015.
[24] S. Maldonado, R. Weber and J. Basak, "Simultaneous Feature Selection and Classification using Kernel-Penalized Support Vector Machines," *Information Sciences*, vol. 181, pp. 115–128, 2011.
[25] O.L. Mangasarian, "Mathematical Programming in Neural Networks," *Journal on Computing*, vol. 5, no. 4, pp.349–360, 1993.
[26] O.L. Mangasarian, "Mathematical Programming in Data Mining," *Data Mining and Knowledge Discovery*, vol. 42, no. 1, pp. 183–201, 1997.
[27] A. McAfee and E. Brynjolfsson, "Big Data: The Management Revolution," *Harvard Business Review*, October, pp. 60–68, 2012.
[28] C.A. Micchelli, "Interpolation ofScattered Data: Distance Matrices and Conditionally Positive Definite Functions," *Constructive Approximation*, vol. 2, pp. 11–22, 1986.
[29] S. Murray, *Interactive Data Visualization for the Web*, O'Reilly Media, 2013.
[30] L. Qi, "Convergence Analysis of Some Algorithms for Solving Nonsmooth Equations," *Mathematics of Operations Research*, vol. 18, pp. 227–244, 1993.
[31] L. Qi and J. Sun, "A Nonsmooth Version of Newtons Method," *Mathematical Programming*, vol. 58, pp. 353–368, 1993.
[32] K. Reda, A. Febretti, A. Knoll, J. Aurisano, J. Leigh, A. Johnson, M.E. Papka and M. Hereld, "Visualizing Large, Heterogeneous Data in Hybrid-Reality Environments," *IEEE Computer Graphics and Applications*, vol. 33, no. 4, pp. 38–48, 2013.
[33] R. Savitha, S. Suresh and N. Sundararajan, "Fast Learning Circular Complex-Valued Extreme Learning Machine (CC-ELM) for Real-Valued Classification Problems," *Information Sciences*, vol. 187, pp. 277–290, 2012.
[34] B. Schölkopf, C.J.C. Burges and A.J. Smola, *Advances in Kernel Methods: Support Vector Learning*, Cambridge, Massachusetts, MIT Press, 1999.
[35] T. Shelton, A. Poorthuis, M. Graham and M. Zook, "Mapping the Data Shadows of Hurricane Sandy: Uncovering the Sociospatial Dimensions of Big Data," *Geoforum*, vol. 52, pp. 167–179, 2014.

[36] H. Tanaka and H. Lee, "Interval Regression Analysis by Quadratic Programming Approach," *IEEE Transactions on Fuzzy Systems*, vol. 6, pp. 473–481, 1998.
[37] A. Unler, A. Murat and R.B. Chinnam, "mr^{2}PSO: A Maximum Relevance Minimum Redundancy Feature Selection Method Based on Swarm Intelligence for Support Vector Machine Classification," *Information Sciences*, vol. 181, pp. 4625–4641, 2011.
[38] S.R. Upadhyaya, "Parallel Approaches to Machine Learning–A Comprehensive Survey," *Journal of Parallel and Distributed Computing*, vol. 73, no. 3, pp. 284–292, 2013.
[39] V.N. Vapnik, *Statistical Learning Theory*, John Wiley and Sons, Inc., New York, 1998.
[40] P. Zikopoulos, C. Eaton and P. Zikopoulos, *Understanding Big Data: Analytics for Enterprise Class Hadoop and Streaming Data*, McGraw-Hill Companies, Inc., 2012.