Industrial Engineering & Management
ISSN: 2169-0316
Wei, Ind Eng Manage 2013, 2:4
DOI: 10.4172/2169-0316.1000e117
Editorial Open Access

Data Mining for Industrial Engineering and Management

Chih-Chiang Wei*
Department of Digit Fashion Design, Toko University, Taiwan

The focus areas of the Industrial Engineering and Management journal include production, logistics, quality, operational research, information systems, technology, communication, industrial economics, regional development, management, organizational behavior, human resources, finance, accounting, marketing, education, training, and professional skills [1]. The aim of this journal is to become a reliable source of information for leaders in the field of industrial engineering management journals research, and to feature a rapid review process [2]. The subject discussed in this paper is data mining for industrial engineering and management.

Knowledge Discovery in Databases

Knowledge Discovery in Databases (KDD) is an iterative process of extracting implicit, previously unknown, and potentially useful knowledge as a production factor from large datasets [3]. It includes data selection, cleaning, integration, transformation, data mining (DM), and reporting. The KDD process consists of steps that are performed before conducting data mining (i.e., selection, pre-processing, and transformation of data), the actual DM, and subsequent steps (i.e., interpretation, and evaluation of results) [4]. DM refers to the specific step of applying one or more statistical, machine-learning, or image-processing algorithms to a particular dataset with the intent to extract useful patterns from the datasets [5]. DM is widely used in market segmentation, customer profiling, fraud detection, retail promotions, and credit risk analysis [6].

Data Mining

With the rapid growth of databases in numerous modern enterprises, DM has become an increasingly valuable data analysis approach. The operations research community has made substantial contributions to this field, particularly by formulating and solving numerous DM problems as optimization problems. In addition, several operations research applications can be addressed using DM methods [7]. In recent years, the data mining field has experienced substantial interest from both academia and industry.

DM problems are typically categorized as association, clustering, classification, and prediction [8]. DM involves various techniques, including statistics, neural networks, decision trees, genetic algorithms, and visualization techniques that have been developed over the years.

Statistics

Regression is one of the most crucial statistical methods applied to science, engineering, economics, and management [9]. Regression is useful for the prediction of the presence or absence of a characteristic or outcome based on values of a set of independent variables that are continuous, categorical, or both. Furthermore, it assumes that measures of dependent variables were independently and randomly sampled, all potentially relevant independent variables are in the model, and all independent variables in the model are relevant [10,11].

Artificial neural networks

Artificial neural networks (ANN) have received intense research focus over the past few years, and are being successfully applied across an extraordinary range of problems and diverse domains, such as finance, medicine, engineering, geology, and physics [12]. ANN is a structure of many neurons, connected in a systematic way. ANNs were created to simulate the nervous system and brain activity. After McCulloch and Pitts [13] established the first neural networks, various ANNs such as feed-forward multi-layer perceptron (MLP), radial basis function (RBF) and self-organizing feature mapping (SOFM) neural networks [14] were developed to solve a wide variety of problems [15]. Neural networks efficiently solve attribute dependency problems [16], but suffer from poor interpretability because it is difficult for humans to construe the logical meaning behind learned weights.

Decision trees

A decision tree is a data mining approach that is often used for classification and prediction. In principle, numerous decision trees can be constructed from any set of attributes. Because certain trees are more accurate than others, determining the optimal tree is computationally infeasible, because of the exponential size of the search space [17]. Several well-known decision trees include ID3 [18], CHAID [19], CART [20], and C5.0 [21] which are greed local search algorithms with top-down constructed trees [22-24]. Tree-derived rules are easy to interpret, capable of coping with noisy data, and can be efficiently generated [25-27].

The fields of DM and KDD have emerged as a new discipline in engineering and computer science to address new opportunities and challenges [28]. Industrial Engineering and Management, with the diverse areas it encompasses, presents unique opportunities for the application of DM and KDD, and for the development of new concepts and techniques in this field. The author encourages the scientific community to submit high-level research and education related articles involving novel DM techniques for industrial engineering and management.

References

1. Patel MH (2012) Editorial for Industrial Engineering & Management Journal. Indust Engg Manag 1:e101.
2. Yalaoui F, Dugardin F, Chehade H (2012) Multi-objective Optimization and Engineering Management. Indust Engg Manag 1:e103.
3. Han J, Kamber M (2001) Data Mining: Concepts and Techniques, San Francisco: Morgan Kaufmann Publishers.
4. Crespo F, Weber R (2005) A methodology for dynamic data mining based on fuzzy clustering. Fuzzy Set Sys 150: 267-284.
5. Bradford JP, Fortes JAB (2001) Characterization and parallelization of decision-tree induction, J Parallel Distr Com 61: 322-349.
6. Wei CC, Hsu NS (2008) Derived operating rules for a reservoir operation

***Corresponding author:** Department of Digit Fashion Design, Toko University, Taiwan, E-mail: d89521007@ntu.edu.tw

Received August 16, 2013; **Accepted** August 18, 2013; **Published** August 23, 2013

Citation: Wei CC (2013) Data Mining for Industrial Engineering and Management. Ind Eng Manage 2: e117. doi: 10.4172/2169-0316.1000e117

system: Comparison of decision trees, neural decision trees and fuzzy decision trees. Water Resour Res 44.

7. Olafsson S, Li X, Wu S (2008) Operations research and data mining. Euro J Oper Res 187: 1429-1448.
8. Chien CF, Chen LF (2008) Data mining to improve personnel selection and enhance human capital: A case study in high-technology industry. Expert Sys App 34: 280-290.
9. Wang S (1999) Nonlinear regression: A hybrid model. Com Oper Res 26: 799-817.
10. Kleinbaum DG (1994) Logistic Regression: A Self-learning. Text. US: Springer-Verlag.
11. Hosmer DW, Lemeshow S (2000) Applied Logistic Regression. US: John Wiley & Sons.
12. Bishop C (1995) Neural Networks for Pattern Recognition. UK: Oxford University Press.
13. McCulloch, WS, Pitts W (1943) A logical calculus of the ideas immanent in nervous activity. B Math Biophy 5: 115-133.
14. Rumelhart DE, Hinton GE, Williams RJ (1986) Learning internal representation by error propagation. Parallel Distr Pro 1: 318-362.
15. Schalkoff RJ (1997) Artificial Neural Networks. New York: McGraw-Hill.
16. Chang FJ, Hu HF, Chen YC (2001) Counter propagation fuzzy neural network for stream-flow reconstruction, Hydrol. Processes 15: 219-232.
17. Tan PN, Steinbach M, Kumar V (2006) Introduction to Data Mining. Boston: Addison Wesley.
18. Quinlan JR (1979) Discovering Rules by Induction from Large Collection of Examples. Edinburg: Edinburg University Press.
19. Kass GV (1980) An exploratory technique for investigating large quantities of categorical data. App Stat 29: 119-127.
20. Breiman L, Friedman JH, Olshen RA, Stone CJ (1984) Classification and Regression Trees. Belmont: Wadsworth.
21. Quinlan JR (1993) C4.5: Programs for Machine Learning. San Mateo: Morgan Kaufmann.
22. Rissanen J, Wax M (1998) Algorithm for Constructing Tree Structured Classifier. US Patent No 4719571.
23. Sreerama KM (1998) Automatic construction of decision trees from data: A multidisciplinary survey. Data Min Knowl Disc 2: 245-389.
24. Kearns M, Mansour Y (1999) On the boosting ability of top-down decision tree learning algorithms. J Comput Syst Sci 58: 109-128.
25. Brodley CE, Utgoff PE (1995) Multivariate decision trees. Mach Learn 19: 45-77.
26. Li XB, Sweigart JR, Teng JTC, Donohue JM, Thombs LA, Wang SM (2003) Multivariate decision trees using linear discriminants and Tabu search. Systems. Man and Cybernet Part A. IEEE Transactions 33: 194-205.
27. Fu Z, Golden BL, Lele S, Raghavan S, Wasil E (2006) Diversification for better classification trees. Comput Oper Res 33: 3185-3202.
28. Triantaphyllou E, Liao TW, Iyengar SS (2002) A focused issue on data mining and knowledge discovery in industrial engineering. Comput Ind Eng 43: 657-659.