

Integration of Rough Set and Affinity Set for Data Mining on the Life-Cycle and Sales of Bio-Product

黃肇偉、黃開義；陳郁文

E-mail: 9606748@mail.dyu.edu.tw

ABSTRACT

In this study, we used Affinity/Guanxi Data Mining the most attributes of the sales at the three stages on the life cycle of Bio product. Since Affinity/Guanxi Set was proposed by Chen and Larbani in 2005, that was just for development. Affinity/Guanxi Set is a time-dependent set theory. First, we interviewed the 10 Bio companies in Taiwan and found the attributes of influencing the sales. In order to focus on the bio product, we choose the most Bio-Product in Taiwan- medicine to be our case. The raw of Bio medicine product in Taiwan is almost imported by Europe and America, R&D is less. So we delete the Development stage and Introduce stage and we just use Growth, Maturity and Decline stage to set the model. our data (include the 9 attributes and sales data) was from a pharmaceutical company in Taipei. Second, we defined the three stages (on the Product Life Cycle by the error of last year sales and this year sales. After separating the data, we used the Rough Set (RS) to find the rules of the three stages. Finally, we must calculate the Hit Rates of the rules and find the k-core. Therefore, we discussed relation between the most attributes and Product Life Cycle. As the result, we find Affinity Set is better method than Rough Set of choosing rules. In the part of Bio medicine product, we got the most important factors on the Bio-Product Life Cycle efficiently. The pharmaceutical company can be referenced by this study.

Keywords : Affinity/Guanxi Set (AS) ; Data Mining ; Product Life Cycle ; Bio medicine product ; Rough Set (RS) ; k-core

Table of Contents

TABLE OF CONTENTS 封面內頁 簽名頁 授權書.....	1	ABSTRACT.....	IV	中文摘要.....	V	誌謝.....	VI	TABLE OF CONTENTS.....	VI	LIST OF TABLE.....	IX	LIST OF FIGURE.....	X	CHAPTER 1. INTRODUCTION.....	1	1.1 MOTIVATION.....	1	1.2 PURPOSE OF THE STUDY.....	2	1.3 THE RESEARCH PROCEDURES.....	2	1.4 THESIS ORGANIZATION.....	3	CHAPTER 2. LITERATURE REVIEW.....	5	2.1 DATA MINING.....	5	2.1.1 Introduction of Data Mining.....	6	2.1.2 Method of Data Mining.....	6	2.2 SET REVIEW.....	6	2.2.1 Fuzzy Set.....	7	2.2.1.1 Basic Definitions.....	8	2.2.1.2 Fuzzy Set Operations.....	9	2.2.2 Rough Set	11	2.2.2.1 The equivalence relation.....	11	2.2.2.2 Set Approximation.....	12	2.2.2.3 Rough Membership Function.....	12	2.3 BIOTECHNOLOGY.....	13	2.3.1 The definition of Biotechnology.....	14	2.3.2 History of Biotechnology.....	14	2.3.3 Bio Industry market in Taiwan.....	15	2.3.4 Pharmaceutical company.....	16	2.4 PRODUCT LIFE CYCLE.....	18	2.4.1 Introduction of Product Life Cycle.....	18	2.4.2 Analysis of Product Life Cycle Model.....	19	2.4.3 The assumption of PLC Model in this study.....	21	2.5 DISCUSSION OF CORRELATIVE PAPER OF DATA MINING.....	22	CHAPTER 3. REDUCTION OF THE AFFINITY SET MODEL.....	25	3.1 BASIC CONCEPTS OF AFFINITY SET.....	25	3.2. INDIRECT AFFINITY.....	29	3.3. OPERATIONS ON AFFINITY SETS.....	30	3.4. POTENTIAL APPLICATIONS.....	31	3.4.1. Forecasting.....	31	3.4.2 Financial State Forecasting.....	35	3.5. AFFINITY SET DEPENDING ON TIME AND OTHER VARIABLES.....	39	CHAPTER 4. RESEARCH METHODOLOGY.....	45	4.1 FRAME OF RESEARCH.....	45	4.2 METHOD OF RESEARCH.....	46	4.3 OBJECT OF RESEARCH.....	46	4.4 RESULT OF THE QUESTIONNAIRE.....	46	4.5 DEFINITION OF THE THREE STAGES ON THE PRODUCT LIFE CYCLE BY AFFINITY SET.....	48	4.6 INTRODUCTION OF ROSETTA SOFTWARE.....	51	CHAPTER 5. DESIGN AND ANALYSIS OF THE EXPERIMENT.....	54	5.1 NUMERICAL EXAMPLE.....	54	5.2. ANALYSIS OF THE RESULT.....	58	CHAPTER 6. CONCLUSION AND RECOMMENDATION.....	60	6.1 CONCLUSION.....	60	6.2 RECOMMENDATION.....	61	REFERENCES.....	63
-------------------------------------	---	---------------	----	-----------	---	---------	----	------------------------	----	--------------------	----	---------------------	---	------------------------------	---	---------------------	---	-------------------------------	---	----------------------------------	---	------------------------------	---	-----------------------------------	---	----------------------	---	--	---	----------------------------------	---	---------------------	---	----------------------	---	--------------------------------	---	-----------------------------------	---	-----------------------	----	---------------------------------------	----	--------------------------------	----	--	----	------------------------	----	--	----	-------------------------------------	----	--	----	-----------------------------------	----	-----------------------------	----	---	----	---	----	--	----	---	----	---	----	---	----	-----------------------------	----	---------------------------------------	----	----------------------------------	----	-------------------------	----	--	----	--	----	--------------------------------------	----	----------------------------	----	-----------------------------	----	-----------------------------	----	--------------------------------------	----	---	----	---	----	---	----	----------------------------	----	----------------------------------	----	---	----	---------------------	----	-------------------------	----	-----------------	----

REFERENCES

- References [1] Chen, Larbani (2006), Developing the Affinity Set (or Guanxi Set) Theory and Its Applications, working paper.
- [2] Larbani, Chen (2007), A fuzzy Set Based Framework for the Concept of Affinity, working paper.
- [3] Haack (1979), Do we need fuzzy logic? International Journal of Man-Machine Studies, Vol.11, pp 437-445.
- [4] Ho (1998), Interpersonal Relationships and Relationship Dominance: An Analysis Based on Methodological Relationism, Asian Journal of Social Psychology, Vol. 1, No. 1, pp. 1-16.
- [5] Hrbacek, Jech (1999), Introduction to Set Theory, Third Edition, Marcel Dekker, Inc., New York.
- [6] Bellman, Zadeh (1970), Decision making in a fuzzy environment, Management Science 17B, 141-164.
- [7] <http://www.ortech-engr.com/fuzzy/reservoir.html> [8] Hwang (1980), Face and Favor: The Chinese Power Game, The American Journal of Sociology, Vol. 92, No. 4, pp. 944-974.
- [9] Khalil (2000), Management of Technology: The Key to Competitiveness and Wealth Creation, McGraw Hill Ltd., New York.
- [10] Luo (2000), Guanxi and Business, World Scientific.
- [11] Pawlak (1982), Rough Sets. International Journal of Computer and Information Science, Vol. 11, pp. 341-356.
- [12] Pawlak (1991), Rough Sets. Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers.
- [13] Zadeh (1965), Fuzzy Sets, Information and control 8 338-353.
- [14] Xiu (2003), proposed Discovery the Effect on Innovation Intellectual Capital to New Product Development Performance: the Case study in Pharmaceutical Industry ", working paper.
- [15] NETMBA Business Knowledge Center www.netmba.com/marketing/product/lifecycle/ [16] Bass (1969), A New Product Growth Model for Consumer Durables.
- [17] Liao (2002) " Study on application of data mining for pharmaceutical market in Taiwan " [18] Liao, Chen, Cheng (2004) " Optimal Investment Decision and Product Life Cycle - A Real Options Approach " , Sun Yat-Sen Management Review [19] Anon : Biotechnology. Report of a joint working party (1980). HMSO, London.
- [20] Biotechnology Weekly Report (2003), BioCentury.
- [21] Biotechnology Industry in Taiwan (2000): DCB, BiotechEast.
- [22] Brealey, Stewart Myers. (1991). Principles of corporation finance. 4th ed. New York: McGraw-Hill [23] Brigham, Gapenski. (1994). Finance Management. 7th ed. Fort Worth: Dryden Press.
- [24] Booz, Hamilton (1982) , Management of New Products for 1980" , New York: Booz, Allen & Hamilton Inc. , p.17-19 [25] Cooper (1983) , Dimensions of Industrial New Product Success and Failure , Journal of Marketing , Vol.34, Summer.
- [26] Kuczmarski (1988) , New Product Development: Steps & Processes , J. Product Innovation Management , p17-29.
- [27] Engel and Blackwell(1982) , Consumer Behavior , New York: Holt , Rinehart & Winston.
- [28] Yip (1989) , Case Study Research: Design and Methods , Revised Ed. , Beverly Hills , CA: Sage.
- [29] Jorquera, Espejo, M.V. Leterlieuw, and G. Acuna (1998), Forecasting ozone daily maximum levels at Santiago, Chile, Atmospheric Environment, 32(20), 3415-3424.
- [30] Kalogirou (2003), Artificial intelligence for the modeling and control of combustion processes: a review, Progress in Energy and Combustion Science, 29, 515-566.
- [31] Kandari, Hawary (2004), Fuzzy short-term electric load forecasting, Electrical Power and Energy Systems, 26, 111-122.
- [32] Karavezyris, Marzi (2002), Application of system dynamics and fuzzy logic to forecasting of municipal solid waste, Mathematics and Computers in Simulation, 60, 149 – 158.
- [33] Karnik, Mendel (1999), Applications of type-2 fuzzy logic systems to forecasting of time-series, Information Sciences, 120, 89-111.
- [34] Kuo (2001), A sales forecasting system based on fuzzy neural network with initial weights generated by genetic algorithm, European Journal of Operational Research, 129, 496-517.
- [35] Kuo, Hwang (2001), An intelligent stock trading decision support system through integration of genetic algorithm based fuzzy neural network and artificial neural network, Fuzzy Sets and Systems, 118, 21-45.
- [36] Kuo, Xue1 (1998), A decision support system for sales forecasting through fuzzy neural networks with asymmetric fuzzy weights, Decision Support Systems, 24, 105 – 126.
- [37] Kuo, Xue2 (1999), Fuzzy neural networks with application to sales forecasting, Fuzzy Sets and Systems, 108, 123-143.
- [38] Kuo, Wang (2002), An intelligent sales forecasting system through integration of artificial neural networks and fuzzy neural networks with fuzzy weight elimination, Neural Networks, 15, 909-925.
- [39] Lee, Chen (2001), Fuzzy regression model with fuzzy input and output data for manpower forecasting, Fuzzy Sets and Systems, 119, 205-213.
- [40] Liao, Tsao (2004), Application of fuzzy neural networks and artificial intelligence for load forecasting, Electric Power Systems Research, 70, 237 – 244.
- [41] Luchetta, Manetti (2003), A real time hydrological forecasting system using a fuzzy clustering approach, Computers & Geosciences, 29, 1111-1117.
- [42] Maier, Lence (2001), Forecasting cyanobacterium Anabaena spp. in the River Murray, South Australia, using B-spline neurofuzzy models,

Ecological Modelling, 146, 85-96.

- [43] Mastorocostas, Bakirtzis (2000), A hybrid fuzzy modeling method for short-term load forecasting, *Mathematics and Computers in Simulation*, 51, 221 – 232.
- [44] Mastorocostas, Petridis (2001), A constrained orthogonal least-squares method for generating TSK fuzzy models: Application to short-term load forecasting, *Fuzzy Sets and Systems*, 118, 215-233.
- [45] Morabito, Versaci (2003), Fuzzy neural identification and forecasting techniques to process experimental urban air pollution data, *Neural Networks*, 16, 493-506.
- [46] Padmakumari, Thiruvengadam (1999), Long term distribution demand forecasting using neuro fuzzy computations, *Electrical Power and Energy Systems*, 21, 315-322.
- [47] PalshiKar (2001), A fuzzy temporal notation and its application to specify fault patterns for diagnosis, *Pattern Recognition Letters*, 22, 381-394.
- [48] Papadakis, Bakirtzis (2003), A load curve based fuzzy modeling technique for short-term load forecasting, *Fuzzy Sets and Systems*, 135, 279-303.
- [49] Peton, Vuillot (2000), Modelling and analysis of ozone episodes, *Environmental Modelling & Software*, 15, 647-652.
- [50] Pokrovskya, Ng (2002), Fuzzy logic approach for description of meteorological impacts on urban air pollution species: a Hong Kong case study, *Computers & Geosciences*, 28, 119-127.
- [51] Profillidis (2000), Econometric and fuzzy models for the forecast of demand in the airport of Rhodes, *Journal of Air Transport Management*, 6, 95-100.
- [52] Ranaweera, Karady (1996), Fuzzy logic for short term load, *Electrical Power and Energy Systems*, 18(4), 25-222.
- [53] Ring, Swan (1979), Product life circle research: a literature review, *Journal of Business Research*, 78(3), 219-242.
- [54] Sakai, Oosumi (1999), Development of a fuzzy sales forecasting system for vending machines, *Computers & Industrial Engineering*, 36, 427-449.
- [55] Saruwatari, Yomita (1995), Forecasting system of irrigation water on paddy field by fuzzy theory, *Agricultural Water Management*, 28, 163-178.
- [56] See, Abrahart (2001), Multi-model data fusion for hydrological forecasting, *Computers & Geosciences*, 27, 987-994.
- [57] Sharaf and Lie (1995), A novel neuro-fuzzy based self-correcting online electric load forecasting model, *Electric Power Systems Research*, 34, 121-125.
- [58] Sisman, Jainc (2004), ANFIS unfolded in time for multivariate time series forecasting, *Neurocomputing*, 61, 139-168.
- [59] Studer, Masulli (1997), Building a neuro-fuzzy system to efficiently forecast chaotic time series, *Nuclear Instruments & Methods in Physics Research*, A389, 264-267.
- [60] Thomassey, Castelain (2005), A short and mean-term automatic forecasting system-application to textile logistics, *European Journal of Operational Research*, 161, 275-284.
- [61] Tsaur (2003), Extrapolating Internet Users in Taiwan by Risk Assessment, *Computers and Mathematics with Applications*, 46, 1725-1734.
- [62] Tseng, Tzeng (2002), A fuzzy seasonal ARIMA model for forecasting, *Fuzzy Sets and Systems*, 126, 367-376.
- [63] Tseng, Yu (1999), Fuzzy Seasonal Time Series for Forecasting the Production Value of the Mechanical Industry in Taiwan, *Technological Forecasting and Social Change*, 60, 263-273.
- [64] Tseng, Yuan (2001), Fuzzy ARIMA model for forecasting the foreign exchange market, *Fuzzy Sets and Systems*, 118, 9-19.
- [65] Tseng, Lin (2005), A quadratic interval logit model for forecasting bankruptcy, *Omega*, 33, 85-91.
- [66] Turkhen, Willson (1995), A fuzzy set model for market share and preference prediction, *European Journal of Operational Research*, 82, 39-52.
- [67] Viett (1965), Exploit the product life cycle, *Harvard Business Review*, 43(9-10), 81-94.
- [68] Xiang, Kieran (2001), A non-linear combination of the forecasts of rainfall-runoff models by first-order Takagi-Sugeno fuzzy system, *Journal of Hydrology*, 245, 196-217.
- [69] Zadeh (1965), Fuzzy sets and systems, In: Fox, J., ed., *System Theory*, 338-353, Polytechnic Press, Brooklyn, NY.