Hybrid Genetic-Fuzzy System Modeling Application in Innovation Management*

Kemal Kilic[1] and Jorge Casillas[2]

[1] FENS, Sabanci University, Tuzla, Istanbul, Turkey
[2] Dept. Computer Science and Artificial Intelligence, Research Center on Information and Communication Technologies (CITIC-UGR), University of Granada, Spain

Abstract. In this research a three staged hybrid genetic-fuzzy systems modeling methodology is developed and applied to an empirical data set in order to determine the hidden fuzzy if-then rules. The empirical data was collected in an earlier study in order to establish the relations among human capital, organizational support and innovativeness. The results demonstrate that the model based on the fuzzy if-then rules outperforms more traditional techniques. Furthermore, the proposed methodology is a valuable tool for successful knowledge management.

Keywords: Knowledge Management, Innovation.

1 Motivation

Knowledge Management (KM) focuses to the generation, sharing, maintenance, refinement and application of knowledge in organizations. The earliest applications of KM are mostly limited to various hardware solutions for data storage and processing. On the other hand, with the advancement of information technologies and information system tools terabytes of data is being gathered by the firms from all sorts of processes and transactions. However, *data* has limited value (if any) unless the patterns hidden in it (*information*) are brought to surface and transformed to capability to act (*knowledge*). Therefore, concepts such as artificial intelligence, decision support systems and data mining enhance the KM capabilities, which is invaluable for the sustainable competitiveness of companies.

Operations Research and Management Science (OR/MS) literature mostly focuses on operational level decision making problems and seems to neglect the strategic level problems due to the vagueness and/or lack of the mathematical formulation of such problems. However, strategic level decisions have significant influence on the competitiveness of the companies. Therefore there is a tremendous need for tools that can be utilized by the senior managers, which guides their

* This work has been supported in part by the *Andalusian Government* (Spain) under grant P07-TIC-3185 and in part by The Scientific and Technological Research Council of Turkey under grant TUBITAK -2219.

J. Casillas et al. (Eds.): Management Intelligent Systems, AISC 171, pp. 25–34.
springerlink.com

strategic level decisions based on knowledge. The first step of developing such tools is generating the relevant knowledge. The upper level managers' knowledge on strategic issues is mostly based on experience and common sense. Therefore, the OR/MS researchers should focus more to strategic level issues and generate scientifically validated knowledge based on objective data and methodologies.

This paper aims to contribute to reduce the existing gap in the literature. For this purpose, with the help of artificial intelligence tools, empirical data collected regarding to a strategic level decision making problem is analyzed. Therefore the research is part of *knowledge generation* process, particularly generation of *explicit knowledge* for a strategic level decision making problem based on objective data. On the other hand, the main research problem, i.e., the determination of the role of human capital and organizational support in the firm level innovation process, is also closely related with KM; hence, the resulting knowledge from the analysis *itself* is highly relevant to the KM literature.

The paper also aims to contribute to the literature in terms of the methodology that is utilized in the data mining phase of such research. Generally speaking there are two different methodological perspectives in data mining process. These are namely the *predictive induction* and the *descriptive induction*. The predictive induction is conducted with the purpose of classification and mostly utilizes supervised learning techniques. On the other hand, the main objective of the descriptive induction is identification of *interesting patterns* and mostly utilizes unsupervised learning techniques (Casillas and Martínez-López, 2009).

Both of the perspectives are sound and valid. Therefore, the *most appropriate* perspective that should be utilized in the analysis depends totally on the application itself. In this research, the main objective is generation of explicit knowledge that can be utilized by the senior managers. Basically, the developed methodology should yield "valuable" information for the decision makers. Valuable information implies that it should be understandable by the managers (so that it can easily be transformed to knowledge), it should address various different subgroups so that the resulting knowledge spans a wide range of the universe and at the same time the information should represent a model that *fits* to the empirical data (i.e., have an acceptable predictive accuracy). Hence, the application on hand, requires an approach which synthesis the two perspectives in order to employ the best from the both approaches. Therefore, in this research, a novel methodology which resembles both *descriptive* (Martínez-López and Casillas, 2009) and *predictive* (Casillas and Martínez-López, 2009) induction is developed.

Among various different techniques in the fields of data mining and machine learning, Fuzzy System Modeling (FSM) emerges as a promising tool that can be utilized in the analysis. Hence, a hybrid genetic-fuzzy system modeling methodology which utilizes the multiobjective genetic fuzzy system approach (Casillas and Martínez-López, 2009; Martínez-López and Casillas, 2009) is developed. Note that, the interpretability of the fuzzy if-then rules, which are usually characterized in natural language, is what sets apart FSM among other candidates and makes it the most relevant choice in this paper. On the other hand, by incorporating the *predictive accuracy* as one of the objective functions during the rule selection process, its ability to represent the empirical data accurately is also addressed.

To sum up, in this paper a novel hybrid genetic-fuzzy system modeling methodology is developed which utilizes multiobjective genetic fuzzy systems applied to a strategic level decision making problem namely innovation management in order to generate knowledge that might be useful for the senior managers. The developed framework is not limited to the particular research problem but is also applicable to various strategic level management problems as well.

In the next section, relevant literature regarding to the organizational support, human capital and innovation performance as well as the problem statement will be presented. In Section 3, the methodology that will be utilized in the research will be provided. Later a brief summary of the data set will be provided in Section 4. The details of the experimental analysis and the results will be covered in Section 5. The paper will be concluded with some concluding results.

2 Organizational Support, Human Capital and Innovativeness

Innovation can be described as the value adding changes in business processes, services, products, marketing or the ways that the works are organized in a company. Schumpeter (1934) differentiated between five different types of innovation: new products, new methods of production, new sources of supply, the exploitation of new markets, and new ways to organize business. Literature on innovation management demonstrates that companies should manage their innovation performance carefully in order to stay competitive. On the other hand, the *innovativeness* of the company refers to the capability of making innovations or the degree of success of the innovation management performance.

Various factors are shown to be influencing the firm level innovativeness. Among these *organizational support* enhances the innovativeness particularly at the individual employee level. Organizational support can be shaped by some managerial arrangements, such as *work discretion*, *rewarding systems, management support* for generation of new ideas, allocation of *time availability*, and *tolerance for failures* in creative undertakings and risky innovation projects (Kuratko et al., 1990; Hornsby et al., 2002; Alpkan et al., 2010).

Work discretion, that is to say the ability to take initiative in decision making is demonstrated to enhance the innovativeness and overall performance of the companies(e.g. Alpkan et al., 2007). On the other hand, high level of trust among the employees in the reward system of their company also positively influences their commitment to innovativeness (Morrison and Robinson, 1997). Management support, i.e., the encouragement of new idea generation and development, positively influence a firms' entrepreneurial behavior and enhance potential intrapreneurs' willingness to innovate (Stevenson and Jarillo, 1990). Free time availability is demonstrated as another critical factor for the employers both daily routines and intrapreneurial ideas and activities (e.g., Fry, 1987) since most of the enthusiastic intrapreneurs make their pioneering steps to actualize their idealized projects in their spare times (Ende et al., 2003). Besides, if the employees feel free from any punishment, adverse criticism, or loss of support in case of failure of their projects or ideas, then their commitment to innovative attempts will be increased (e.g., Morrison and Robinson, 1997; Chandler et al., 2000).

Intellectual capital is the total stocks of all kinds of intangible assets, knowledge, capabilities, and relationships, etc, at employee level and organization level within a company. It is examined in the literature under three subgroups; namely, human, social and organizational capital. The *human capital* is the sum of knowledge and skills that can be improved especially by education and work experience of the employees of an organization (Dakhli and De Clercq, 2004). Hall and Mairesse (2006) states that a great deal of the knowledge created by firm activities is embedded in the human capital. Also, Cohen and Levinthal (1989) suggested that the human capital of a firm is crucial in terms of the innovativeness, due to its ability to obtain and make use of the outcomes of other firms' R&D activities. Human capital also enhances the organizational competencies of the firms by increasing the returns from the innovations and reducing the risks (Hayton 2005). Hence, the human capital not only has direct effect on innovativeness but is also a precious resource that may act as a moderator in the relationship of organizational support and innovativeness (Alpkan et al., 2010).

Alpkan et al. (2010) utilizes an empirical data set in order to analyze the relationship between the organizational support factors, human capital and innovativeness as well as the hypothesized moderator role of human capital. By using multiple linear regression analysis they managed to show that certain dimensions of organizational support factors such as the management support and tolerance for failures have positive influence on innovativeness but the relation of the other three organizational support factors (i.e., time availability, work discretion and reward system) and innovativeness was not supported. On the other hand, human capital is also demonstrated to have positive influence on innovativeness however regarding to the moderation role of human capital, only limited knowledge was attained. The reason for this was based on the limitations associated with the capabilities of the multiple linear regression analysis. Therefore, the knowledge attained as the result of the research was quite limited and left space for further analysis. One of the goals of this paper is utilizing the fuzzy system modeling approach in the same data set in order to attain further valuable knowledge for the decision makers.

3 Hybrid Genetic-Fuzzy System Modeling

The hybrid genetic-fuzzy system modeling methodology that is utilized in the research consists of three stages. The first stage is the identification of a large set of individual fuzzy if-then rules that might be of interest, based on their *confidence* (i.e, a measure of accuracy of the rule) and *support* (i.e., a measure of representativeness of the rule). Hence the problem dealt in this stage is basically a multi objective (accuracy vs. representativeness) decision making problem in which a set of fuzzy if-then rules are obtained as the result. These fuzzy if-then rules should have better scores from the other candidates that were not selected through out the process, in terms of the two conflicting objectives, i.e., their confidence and their support. Therefore, for this stage the multi objective genetic fuzzy system proposed in (Martínez-López and Casillas, 2009) is utilized.

Briefly speaking the genetic algorithm utilizes a gene pool which consists of chromosomes that represents a fuzzy if-then rule regarding to the relationship among the organizational support, human capital and innovativeness. In the model, the innovativeness is set to be the *consequence* of the fuzzy if-then rules and the human capital and organizational support factors (five of them) are set to be the antecedents of the fuzzy if-then rules. Note that, two fuzzy sets are assumed to represent each one of the antecedents (low - L and high - H). On the other hand five crisp values are assumed to represent the consequence in the if-then rules (due to the fact that the data set collected utilized five-point Likert scale as we will discuss in the next section). Therefore the size (i.e., width) of the chromosomes in the gene pool is equal to 17 ($2 + 5 \cdot 2 + 5$; L and H for human capital + L and H for five organizational support factors + 1..5 as the crisp output score for the antecedent, that is to say innovativeness).

The basic genetic operators such as selection, crossover and mutation are utilized as suggested in Martínez-López and Casillas (2009) in the algorithm. Furthermore, the support and the confidence of the fuzzy if-then rules are also determined as suggested in Martínez-López and Casillas (2009). On the other hand, the fitness function scores of the chromosomes are determined based on their domination rank and crowding distance values obtained from their support and confidence.

The second stage of the methodology is merely the elimination of the inconsistent rules among the set of fuzzy if-then rules that are obtained after the first stage. Note that the resulting fuzzy rule set after the first stage might include rules that are redundant and/or contradicting with others in the rule set. An example of an inconsistent rule is a rule which has antecedents that are subsumed by another rule but the output of the subsumed rule is different from the other one. In such cases the rules that has higher confidence is preserved in the set of fuzzy if-then rules and the other one is eliminated. Hence at the end of the second stage, the resulting set of fuzzy rules of stage one reduces in size and only includes consistent fuzzy if-then rules in order to be utilized in the final stage.

The third stage of the methodology is basically, selection of the best set of fuzzy rules among the set of *consistent* fuzzy rules obtained as the result of stage two. This process is also a multi criteria decision making problem since there are two objectives, namely maximization of the *prediction accuracy* (i.e., minimization of the root-mean-square error, RMSE) and minimization of *the number of rules being used* in the set. The prediction accuracy is important since it is kind of a measure of the *goodness-of-fit* of the model to the empirical data. On the other hand the number of rules being used should also be minimized simultaneously in order to assess the knowledge hidden in the data set (i.e., in order to enhance the interpretability and descriptiveness of the attained fuzzy if-then rules). Therefore, for the third stage, the algorithm that is based on the multi objective fuzzy-genetic system which was utilized in the first stage is adopted to the particular problem.

This time the gene pool consists of chromosomes (with size equal to the *cardinality* of the consistent fuzzy rule set which is obtained after the second phase) that represents a candidate set of fuzzy rules (1 represents that the rule is used and 0 represents it is not part of the final rule base). The fitness function is also similar to the fitness function used in stage one, i.e., the domination rank and the crowding distance values obtained from the *prediction accuracy* scores and the *number of rules being used* in the particular chromosome. Again the selection, crossover and mutation genetic operators are utilized as described in Martínez-López and Casillas (2009). Note that the prediction accuracy of each candidate set of fuzzy if-then rules are assessed by means of the training RMSE. The RMSE calculation is conducted by the methodology that was suggested by Casillas and Martínez-López (2009) for the predictive approach.

4 Data Collection

A questionnaire is developed for the empirical survey (Alpkan et al., 2010). In order to measure the human capital five criteria are constructed which were inspired from Subramaniam and Youndt (2005). Similarly organizational support measures were also adapted from several criteria in the Operations Management literature based on previous studies of Kuratko et al. (1990) and Hornsby et al. (2002). On the other hand the innovative performance is measured by means of a scale consisting of the items adapted from the earlier studies of Antoncic and Hisrich (2001), Neely and Hii (1998) and Hagedoorn and Cloodt (2003). All items were measured on a five point Likert scale as suggested in the literature. After the questionnaire was developed, the initial survey draft was discussed with various firms' executives and it was pre-tested through 10 pilot interviews to ensure that the wording, format and sequencing of questions are appropriate. Data was collected over a 7-month period using a self-administered questionnaire distributed to firms' upper level managers operating in manufacturing sectors in the Northern Marmara region in Turkey. A sample of 1,672 manufacturing firms was obtained by selecting randomly from various databases. Afterwards, the questionnaire was applied through a hybrid system of mail surveys and face-to-face interviews. Out of the sample of 1,672 firms, 184 complete responses were obtained resulting in 11% return rate. The data was later controlled with t-test procedure for non-respondent bias and no significant difference ($p \leq 0.05$) was found between the interview and mailing data sets' responses both in terms of the questionnaire items and constructs. Moreover, the issue of Common Method Variance was also attended.

Exploratory factor analysis (EFA) with varimax rotation and confirmatory factor analysis (CFA) to explore and confirm the latent factor structure of the innovative performance, human capital and organizational support factors' scales was conducted. The factor analyses (EFA and CFA) revealed that the hypothesized seven factors were sufficiently valid and reliable (with Cronbach's Alpha value ranging from 0.72 and 0.92 for the constructs). The seven constructs to be used in the analysis were namely, Human Capital, Performance Based Reward System, Management Support for Idea Generation, Tolerance for Risk Taking, Work Discretion, Allocation of Free Time (the latter five factors constitute the components of Organizational Support) and Innovativeness. The former six factors are treated as the inputs (antecedents of the fuzzy if-then rule) in the model where as the Innovativeness is treated as the output (the consequent).

5 Experimental Analysis and Results

A typical GA based algorithm requires at least four parameters to be tuned, namely, *Gene Pool Size* (GPS), *Number of Iterations* (NoI), *Mutation Probability* (MP) and *Crossover Probability* (CP). Since the multi objective genetic fuzzy systems algorithm is employed at two different stages 2·4=8 parameters were considered. For each one of the parameters low and high values are assigned (after some test runs) and an experimental analysis is conducted for parameter tuning. The resulting parameter values used in the analysis is GPS-1=100, NoI-1 = 20, MP-1 = 0.2, CP-1 = 0.6, GPS-2 = 100, NoI-2=40, MP-2=0.2 and CP-2 = 0.9.

The final obtained model consists of seven fuzzy if-then rules and yields a training RMSE equal to 0.388. A similar analysis with a multiple linear regression is conducted and calculated the RMSE in the same manner. The RMSE of the training experiments for the multiple linear regression was determined as 0.455 which suggests that the hybrid genetic-fuzzy system algorithm models the relations in the data better than the MLR (as expected).

Since the Gene Pool Size of the first stage for the tuned parameter set was equal to 100, there were 100 individual fuzzy if-then rules at the end of the first stage. Figure 1, depicts the confidence vs. support degrees of the resulting 100 fuzzy if then rules. Note that from the figure, one can realize how the two objectives, namely the domination rank vs. the crowding distance results.

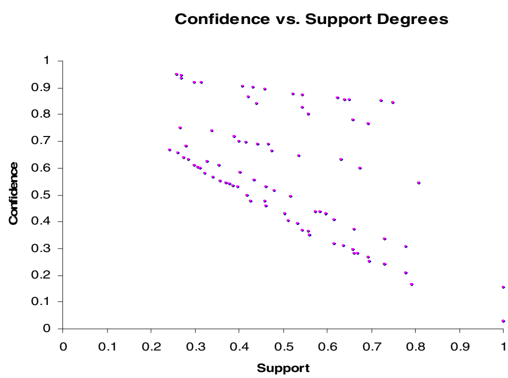

Fig. 1 Confidence vs. Support degrees of the resulting fuzzy if-then rules after the first stage

Among the 100 fuzzy if-then rules, 59 of them were determined to be either subsumed or contradicting with other rules, hence were eliminated during the second stage. Therefore, 41 relevant fuzzy if then rules were used in the third stage of the application. The multi objective genetic–fuzzy system algorithm utilized in the third stage, determined a set of seven rules which yields both good predictive accuracy (so that the empirical data is represented better with the model) and low number of rules (in order to enhance the descriptiveness of the fuzzy sets).

The corresponding seven fuzzy if-then rules that are obtained as the result of the third stage of the methodology and the associated confidence (*Conf.)* and the support (*Sup.*) levels of each fuzzy if-then rules are depicted in Figure 2. Note that Management Support (MS), Tolerance for Risk Taking (RT), Work Discretion (WD), Reward System (RW), Time Availability (TA) and Human Capital (HC) are the antecedents of the fuzzy if-then rules and are represented with two fuzzy sets, i.e., Low (L) and High (H). On the other hand, the consequence is the Innovativeness (I) and represented with five singleton results in which Very Low (VL) refers to a value of 1 and Very High (VH) refers to a value of 5 and the rest accordingly (i.e., Low (L), Medium (M) and High (H), 2, 3 and 4 respectively).

	MS	RT	WD	RW	TA	HC	I	Conf.	Sup.
Rule 1	H	L	H				VH	0.67	0.48
Rule 2	H						VH	0.54	0.81
Rule 3		L	H			L	M	0.87	0.42
Rule 4	L		H		H		M	0.92	0.30
Rule 5		L	L	L		L	VL	0.57	0.34
Rule 6		H	L	H		H	H	0.90	0.46
Rule 7	H	H	L	H			VH	0.69	0.47

Fig. 2 The resulting seven fuzzy if-then rules and associated confidence and support degrees. Training RMSE = 0.388.

The resulting fuzzy if-then rules were parallel with the results of Alpkan et al. (2010) in the sense that these rules also indicated that the Management Support, Tolerance for Risk Taking and Human Capital was positively influencing the innovativeness. On the other hand, the resulting fuzzy rules also demonstrates that the Work Discretion in fact negatively influencing the innovativeness which was not a result of Alpkan et al. but has support in the literature. Note that the Alpkan et al. analysis were merely based on MLR analysis and the utilized hybrid genetic-fuzzy system modeling methodology had a better *goodness-of-fit* (in terms of the prediction accuracy) to the empirical data. Furthermore the resulting fuzzy if-then rules also reveal the lack of the moderator role of human capital on the relation with the organizational support and innovativeness which was hypothesized but couldn't be demonstrated in Alpkan et al. study as well. Particularly Rule 6 demonstrates that whenever the organizational support is high, having human capital high as well not necessarily boosts the innovativeness. Therefore, indicating to a possible *substitute* relation between the human capital and organizational support rather than a *complementary* relation which might suggest synergy among the concepts.

The rules with higher confidence degrees (such as the rules 3, 4 and 7) apparently bear more accurate explicit knowledge in the context. On the other hand, the rules with relatively higher support degrees relate to knowledge on the combinations that are more commonly observed. However, the significance of the proposed methodology lies in the fact that, other rules might be even more interesting instead. In this example Rule 6 (as described above) as well as Rule 5 and Rule 1 reveals highly interesting knowledge for the senior managers.

6 Concluding Remarks

The three staged hybrid genetic-fuzzy system modeling methodology that was applied to a strategic level decision making problem in the context of innovation management demonstrated the strength of the descriptive fuzzy if-then rules in terms of explicit knowledge generation. Furthermore, the better prediction accuracy also hints the ability of fuzzy system modeling to model highly complex and nonlinear systems. The resulting fuzzy if-then rules were highly interesting and revealing and enhance the understanding of the complex interrelations in the problem. Therefore, the developed methodology might be a valuable tool and serve as an engine of a decision support system which might allow the upper level decision makers to make more informed and better decisions.

References

Alpkan, L., Yilmaz, C., Kaya, N.: Market Orientation and Planning Flexibility in SMEs: Performance Implications and an Empirical Investigation. International Small Business Journal 25(2), 152–172 (2007)

Alpkan, L., Bulut, C., Gunday, G., Ulusoy, G., Kilic, K.: Organizational Support for Intrapreneurship and its Interaction with Human Capital to Enhance Innovative Performance. Management Decision 48(5), 732–755 (2010)

Antoncic, B., Hisrich, R.D.: Intrapreneurship: Construct refinement and cross-cultural validation. Journal of Business Venturing 16, 495–527 (2001)

Casillas, J., Martínez-López, F.J.: Mining uncertain data with multiobjective genetic fuzzy systems to be applied in consumer behaviour modeling. Expert Systems with Applications 36, 1645–1659 (2009)

Chandler, G.N., Keller, C., Lyon, D.W.: Unraveling the determinants and consequences of an innovation-supportive organizational culture. Entrepreneurship: Theory and Practice 25(1), 59–76 (2000)

Cohen, W.M., Levinthal, D.A.: Innovation and learning, the two faces of R&D. Economic Journal 99, 569–596 (1989)

Dakhli, M., De Clercq, D.: Human capital, social capital, and innovation: A multicountry study. Entrepreneurship & Regional Development 16, 107–128 (2004)

Ende, J.V.D., Wijnberg, N., Vogels, R., Kerstens, M.: Organizing innovative projects to interact with market dynamics: A coevolutionary approach. European Management Journal 21(3), 273–284 (2003)

Fry, A.S.: The post it note: An intrapreneurial success. SAM Advanced Management Journal 52(3), 4–9 (1987)

Hagedoorn, J., Cloodt, M.: Measuring innovative performance: Is there an advantage in using multiple indicators? Research Policy 32, 1365–1379 (2003)

Hall, B.H., Mairesse, J.: Empirical studies of innovation in the knowledge driven economy. United Nations University Working Paper, 28 (2006)

Hayton, J.C., Zahra, S.A.: Venture team human capital and absorptive capacity in high technology new ventures. International Journal of Technology Management 31(3-4), 256–274 (2005)
Hornsby, J.S., Kuratko, D.F., Zahra, S.A.: Middle managers' perception of the internal environment for corporate entrepreneurship: Assessing a measurement scale. Journal of Business Venturing 17, 253–273 (2002)
Kuratko, D.F., Montagno, R.V., Hornsby, J.S.: Developing an intrapreneurial assessment instrument for an effective corporate entrepreneurship. Strategic Management Journal 11(5), 49–58 (1990)
Martínez-López, F.J., Casillas, J.: Marketing Intelligent Systems for consumer behaviour modelling by a descriptive induction approach based on Genetic Fuzzy Systems. Industrial Marketing Management 38, 714–731 (2009)
Morrison, E.W., Robinson, S.L.: When employees feel betrayed: A model of how psychological contract violation develops. Academy of Management Review 22(1), 226–256 (1997)
Neely, A., Hii, J.: Innovation and Business Performance: A Literature Review. The Judge Institute of Management Studies - University of Cambridge, Cambridge (1998)
Schumpeter, J.A.: The Theory of Economic Development. An Inquiry into Profits, Capital, Credit, Interest, and The Business Cycle. Harvard University Press, Cambridge (1934)
Stevenson, H.H., Jarillo, C.J.: A paradigm of entrepreneurship: Entrepreneurial management. Strategic Management Journal 11(5), 17–27 (1990)
Subramaniam, M., Youndt, M.A.: The influence of intellectual capital on the types of innovative capabilities. Academy of Management Journal 48(3), 450–463 (2005)