Hybrid Model Based on Rough Sets Theory and Fuzzy Cognitive Maps for Decision-Making

Gonzalo Nápoles[1,2], Isel Grau[1], Koen Vanhoof[2], and Rafael Bello[1]

[1] Universidad Central "Marta Abreu" de Las Villas, Santa Clara, Cuba
[2] Hasselt University, Diepenbeek, Belgium
gnapoles@uclv.edu.cu

Abstract. Decision-making could be defined as the process to choose a suitable decision among a set of possible alternatives in a given activity. It is a relevant subject in numerous disciplines such as engineering, psychology, risk analysis, operations research, etc. However, most real-life problems are unstructured in nature, often involving vagueness and uncertainty features. It makes difficult to apply exact models, being necessary to adopt approximate algorithms based on Artificial Intelligence and Soft Computing techniques. In this paper we present a novel decision-making model called Rough Cognitive Networks. It combines the capability of Rough Sets Theory for handling inconsistent patterns, with the modeling and simulation features of Fuzzy Cognitive Maps. Towards the end, we obtain an accurate hybrid model that allows to solve non-trivial continuous, discrete, or mixed-variable decision-making problems.

Keywords: Decision-making, Rough Set Theory, Fuzzy Cognitive Maps.

1 Introduction

In recent years decision-making problems have become an active research area due to their impacts in solving real-world problems. Concisely speaking, decision-making process could be defined as the task of determining and selecting the most adequate action that allows solving a specific problem. This task is supported by the knowledge concerning the problem domain allowing justifying the selected decision. However, the knowledge obtained from experts regularly shows inconsistent patterns that could affect the inference results (e.g. different perception for the same observation).

The Rough Set Theory (RST) is a well-defined technique for handling uncertainty arising from inconsistency [1]. This theory adopts two approximations to describe a set, which are entirely based on the collected data [2] and does not require any further knowledge. Let us assume a decision system $S = (U, A \cup \{d\})$, where U is a non-empty finite set of objects (the universe of discourse), A is a non-empty finite set of attributes, whereas $d \notin A$ is the decision class. Any subset $X \subseteq U$ can be approximated by using two exact sets $B_*X = \{x \in U : [x]_B \subseteq X\}$ and $B^*X = \{x \in U : [x]_B \cap X \neq \emptyset\}$ called lower and upper approximation respectively. In this formulation $[x]_B$ denotes the set of inseparable objects associated to the instance x (equivalence class) using an indiscernibility relation defined by a subset of attributes $B \subseteq A$. The reader may notice that this indiscernibility relation is reflexive, transitive and symmetric.

M. Kryszkiewicz et al. (Eds.): RSEISP 2014, LNAI 8537, pp. 169–178, 2014.

The objects in B_*X are categorically members of X, whereas the objects in B^*X are possibly members of the set X. Notice that this model does not consider any tolerance of errors [3]: if two inseparable objects belong to different classes then the decision system will be inconsistent. The lower and upper approximations divide the universe into three pair-wise disjoint regions: the lower approximation as the positive region, the complement of the upper approximation as the negative region, and the difference between the upper and lower approximations as the boundary region [4]. Being more precise, objects belonging to the positive region $POS(X) = B_*X$ are certainly contained in X, objects that belong to the negative region $NEG(X) = U - B^*X$ are not confidently contained in the set X, whereas the boundary region $BND(X) = B^*X - B_*X$ represents uncertainty about the membership of related objects to the set X.

Such knowledge comprises an opposite knowledge when facing decision-making problems. For example, in reference [5] the author introduces the three-way decisions model. Rules constructed from the three regions are associated to different actions and decisions (see following equations). A positive rule makes a decision of acceptance, a negative rule makes a decision of rejection, and a boundary rule makes a decision of abstaining [4]. Observe that the model interpretation is not so critical in the classical rough set model since it does not involve any uncertainty. In an attempt to overcome this drawback, Wong and Ziarko [6] considered a probabilistic relationship between equivalence classes and X leading to the probabilistic three-way decisions. An object in the probabilistic positive region does not certainly belong to the decision class, but with a high probability. It is important in the probabilistic model, where acceptance and rejection are made with certain levels of tolerance for errors.

- $Des([x]) \rightarrow_P Des(d)$, for $[x] \subseteq POS(d)$
- $Des([x]) \rightarrow_B Des(d)$, for $[x] \subseteq BND(d)$
- $Des([x]) \rightarrow_N Des(d)$, for $[x] \subseteq NEG(d)$

The probabilistic three-way decisions showed to be superior with respect to the original algorithm [7], however, such models are mainly oriented to discrete decision-making problems. In this paper we present a novel hybrid model that combines three-way decisions rules, with the simulation aptitude of Fuzzy Cognitive Maps [8] using a sigmoid threshold function. This hybrid model not just allows to solve mixed-attribute or continuous problems, but also provides accurate inferences. The main idea consists in replacing the equivalence classes by similarity classes to define positive, negative, and boundaries regions. After that, we build a Sigmoid Fuzzy Cognitive Maps (which are a kind of recurrent neural network for modeling and simulation) using computed regions and the domain knowledge. Finally, a recurrent inference process is triggered allowing to the map to converge to a desired decision.

The rest of the paper is organized as follows: in following Section 2 the theoretical background of FCM is described. Here we point out some aspects concerning the map inference process using continuous threshold functions. In Section 3 we introduce the proposed hybrid model consisting in three main steps: (i) the computation of positive, negative and boundary regions, (ii) the construction of the map topology, and (iii) the map exploitation using the similarity class of the target instance. Section 4 provides numerical simulations illustrating the behavior of our algorithm. Finally, conclusions and further research aspects are discussed in Section 5.

2 Fuzzy Cognitive Maps

Fuzzy Cognitive Maps (FCM) are a suitable knowledge-based tool for modeling and simulation [9]. From a connectionist perspective, FCM are recurrent networks with learning capabilities, consisting of nodes and weighted arcs. Nodes are equivalent to neurons in connectionist models and represent variables, entities or objects; whereas weights associated to connections denote the *causality* among such nodes. Each link takes values in the range $[-1,1]$, denoting the causality degree between two concepts as a result of the quantification of a fuzzy linguistic variable, which is often assigned by experts during the modeling phase [10]. The activation value of neurons is also fuzzy in nature and regularly takes values in the range $[0,1]$. Therefore, the higher the activation value of a neuron, the stronger its influence over the investigated system, offering to decision-makers an overall picture of the systems behavior.

Without loss of generality, a FCM can be defined using a 4-tuple (C, W, A, f) where $C = \{C_1, C_2, C_3, \dots, C_M\}$ is a set of M neurons, $W: (C_i, C_l) \to w_{il}$ is a function which associates a causal value $w_{il} \in [-1,1]$ to each pair of nodes (C_i, C_l), denoting the weight of the directed edge from C_i to C_l. The weigh matrix $W_{M\times M}$ gathers the system causality which is often determined by experts, although may be computed using a learning algorithm. Similarly, $A: (C_i) \to A_i$ is a function that associates the activation degree $A_i \in \mathbb{R}$ to each concept C_i at the moment t $(t = 1,2, \dots, T)$. Finally, a transformation function $f: \mathbb{R} \to [0,1]$ is used to keep the neuron's activation value in the interval $[0,1]$. Following Equation (1) portrays the inference mechanism using the vector A^0 as the initial configuration. This inference stage is iteratively repeated until a hidden pattern or a maximum number of iterations T is reached.

$$A_i^{t+1} = f\left(\phi_1 \sum_{j=1}^{M} w_{ji} A_j^t + \phi_2 w_{ii} A_i^t\right), i \neq j \tag{1}$$

In the above equation ϕ_1 represents the influence from the interconnected concepts in the configuration of the new activation value, whereas ϕ_2 regulates the contribution of the neuron memory over its own state. In all experiments conducted in this paper we use $\phi_1 = 0.95$ and $\phi_2 = 1 - \phi_1$ since the new evidence is often desirable.

The most used threshold functions are: the bivalent function, the trivalent function, and the sigmoid variants. It should be stated that authors will be focused on Sigmoid FCM, instead of discrete ones. It is motivated by the benchmarking analysis discussed in reference [11] where results revealed that the sigmoid function outperformed the other functions by the same decision model. Therefore, the proper selection of this threshold function may be crucial for the system behavior. From [12] some important observations were concluded and summarized as follows:

- Binary and trivalent FCM cannot represent the degree of an increase or a decrease of a concept. Such discrete maps always converge to a fixed-point attractor or limit cycle since FCM are deterministic models.
- Sigmoid FCM, by allowing neuron's activation level, can also represent the neuron's activation degree. They are suitable for qualitative and quantitative tasks, however, may additionally show chaotic behaviors.

3 Rough Cognitive Networks

In this section we introduce a hybrid model for addressing decision-making problems called Rough Cognitive Networks (RCN). It combines the ability of RST for handling uncertainty and the simulation strength of FCM. The aim of this model is the mapping of an input vector to a feasible decision, using the knowledge obtained from historical data. Let us consider a set of decisions $D = \{d_1, \dots, d_k, \dots, d_n\}$ for some decision task, a decision system $DS = (U, A, \{d\})$ where problem attributes are mainly continuous, and an unlabeled problem instance O_i. Next steps describe how to design a RCN which is capable of computing the most fitting decision for the new instance O_i by ranking the activation value of decision concepts (i.e. map neurons).

3.1 Determining Positive, Negative and Boundary Regions

The first step of our proposal is oriented to determine positive, negative and boundary regions. It should be stated that we need to use weaker inseparability relations among objects in the universe U, since we assume decision-making problems with continuous or mixed variables. It could be achieved by extending the concept of inseparability, so that they are grouped together in the same class of not identical objects, according to a similarity relationship R. Hence, by replacing the equivalence relation with a weaker binary relation, an extension of the classical RST approach is achieved [2].

Equations (2) and (3) summarize how to compute lower and upper approximations respectively by using this scheme, where $R'(x)$ denotes the similarity class associated to the object x (that is, the set of objects which are similar to the instance x according to some similarity relation R). It suggests that an object can simultaneously belong to different similarity classes, so the covering induced by $R'(x)$ over U is not necessarily a partition [13]. Being more explicit, similarity relations do not provoke a partition of the universe U, but rather generate classes of similarity.

$$B_*X = \{x \in U : R'(x) \subseteq X\} \tag{2}$$

$$B^*X = \bigcup_{x \in X} R'(x) \tag{3}$$

While constructing an equivalence relation is trivial, constructing the apt similarity relation for a problem could be more complex. What is more, the global performance of our model will be reliant on the quality of such similarity relation, however, in the literature several methods for facing this challenge have been proposed. For example, Filiberto et al. [14] describe an optimization procedure for building accurate similarity relations using a population-based metaheuristic.

Another aspect to be considered when designing a similarity relation is the opposite selection of the similarity (or distance) function. It is used for measuring the similarity (or difference) degree between two objects. In reference [15] the authors widely study the properties of several distance functions, where problem attributes are grouped into two large groups: continuous or discrete (both nominal and ordinal). This subject will be detailed in the Section 5, during numerical simulations.

3.2 Designing the Map Topology

During this step, positive, negative and boundary regions are denoted as map neurons denoting input variables of the system. Following a similar reasoning, we use $|D|$ new concepts for measuring the activation degree of each decision. It should be mentioned that a boundaries region will be considered or not (this point will be clarified in next sub-sections). Afterwards, once concepts have been defined, we establish connections among all map neurons where causal values are computed as follows:

- $R_1: if\, C_i\ is\ P_k\ and\ C_l\ is\ d_k\ then\ w_{il} = 1.0$
- $R_2: if\, C_i\ is\ P_k\ and\ C_l\ is\ d_{v \neq k}\ then\ w_{il} = -1.0$
- $R_3: if\, C_i\ is\ P_k\ and\ C_l\ is\ P_{v \neq k}\ then\ w_{il} = -1.0$
- $R_4: if\, C_i\ is\ N_k\ and\ C_l\ is\ d_k\ then\ w_{il} = -1.0$

In rules detailed above C_i and C_l denote two neurons, P_k and N_k are the positive and negative region related to the kth decision respectively, whereas w_{il} denote the causal weight between the cause C_i and the effect C_l. If the positive region P_k is activated, then the FCM stimulates the kth decision node since we confidently know that objects belonging to the positive region will be categorically members of X_k. On the contrary, if the negative region N_k is activated, then the map will inhibit the corresponding decision, but we cannot conclude about other decisions.

Boundary regions usually report uncertain information about the acceptance of the investigated decision, however, an unlabeled object $O_i \in BND(X_k)$ could be correctly associated to decision d_k. Being more specific, let us suppose a problem having three decisions where $O_i \in BND(X_1)$, $O_i \in BND(X_2)$ and $O_i \notin BND(X_3)$. It means that the instance could be labeled as d_1 or d_2 as well, but it provides no evidence supporting decision class d_3. Encouraged by this remark we include another rule where further knowledge about boundary regions is considered.

- $R_5: if\, C_i\ is\ B_k\ and\ C_l\ is\ d_v\ and\ (BND(X_k) \cap BND(X_v) \neq \emptyset)\ then\ w_{il} = 0.5$

Observe that rules $R_1 - R_4$ are independent of the upper and lower approximations, while the rule R_5 requires information about boundary regions. Hence, the final map could have at most $4|D|$ neurons and $3|D|(1 + |D|)$ causal links since the number of boundary concepts will depend on the upper and lower approximations.

3.3 Inferring the Most Fitting Decision

The final phase of this hybrid model is related to the FCM exploitation, so we need to compute the activation value of input concepts (neurons that denote positive, negative and boundary regions). Being more explicit, the excitation vector will be calculated using objects belonging to the similarity class $R'(O_i)$ and their relation to each region. For example, let us suppose that $|POS(X_1)| = 20$, $|R'(O_i)| = 10$, whereas the number of similar objects that belong to the positive region is given by the following expression: $|R'(O_i) \cap POS(X_k)| = 7$. Then the activation degree of the positive neuron associated to the first decision will be $7/20 = 0.35$. Following three rules generalize this scheme for all map concepts (regions), thus complementing the proposal.

- $R_6: if\, C_i\ is\ P_k\ then\ A_i^0 = \frac{|Y|}{|POS(X_k)|}\ where\ Y = R'(O_i) \cap POS(X_k)$
- $R_7: if\, C_i\ is\ N_k\ then\ A_i^0 = \frac{|Y|}{|NEG(X_k)|}\ where\ Y = R'(O_i) \cap NEG(X_k)$
- $R_8: if\, C_i\ is\ B_k\ then\ A_i^0 = \frac{|Y|}{|BND(X_k)|}\ where\ Y = R'(O_i) \cap BND(X_k)$

Figure 1 illustrates the FCM resulting from a decision-making problem having two decisions, assuming inconsistencies on the information system. Notice that each input neuron has a self-reinforcement connection with causal weight $w_{ii} = 1$ which partially preserves the initial knowledge during the simulation (exploitation) stage.

The main reason for adopting FCM as inference mechanism is that they can handle incomplete or conflicting information. Besides, decision-making problems are usually characterized by various concepts or facts interrelated in complex ways, so the system feedback plays a prominent role by efficiently propagating causal influences in non-trivial pathways. Formulating a precise mathematical model for such systems may be difficult or even impossible due to lack of numerical data, its unstructured nature, and dependence on imprecise verbal expressions. Observe that the performance of FCM is dependent on the initial weight setting and architecture [16], but our model provides a general framework for facing problems having different features.

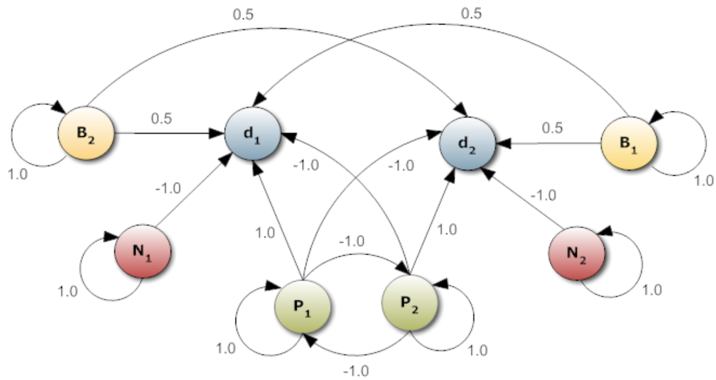

Fig. 1. Resultant FCM model for decision-making problems for two decisions

It should be stated that our algorithm only requires the estimation of a similarity threshold, which is used for building lower and upper approximations. It means that we need to build precise similarity relations to ensure high-quality results. Being more explicit, if this similarity value is excessively small then positive regions will be small as well, leading to poor excitation of related neurons. It could be essential to select the most fitting decision when a new scenario is observed: the more active the positive region, the more desirable the decision (although the algorithm will compute the final decision taking into account all the evidence). If the similarity threshold is too large then boundary regions will be also large, increasing the global uncertainty during the model inference phase, decreasing the overall algorithm performance.

4 Numerical Simulations

In the present section we study the behavior of the proposed RCN approach by using both synthetic and real-life data. In all simulations we adopt a Sigmoid RCN resulting in a FCM having a sigmoid threshold function. The sigmoid function uses a constant parameter $\lambda > 0$ to adjust its inclination. In this paper we use $\lambda = 1$, because this value showed best results in previous studies [16]. During synthetic simulations a decision-making problem having three outcomes d_1, d_2 or d_3 is assumed. With this purpose in mind we evaluate the system response for the following scenarios:

i. the set $R'(O_i)$ activates a single positive region
ii. the set $R'(O_i)$ activates two positive regions
iii. the set $R'(O_i)$ activates boundary regions

To reach the first case we use the excitation values $P_1 = 0.23$ and $N_2 = N_3 = 0.12$ leading to the output (1,0,0) which is the desired solution. It should be stated that we know the expert preference for each target object O_i in advance, but it is only used for measuring the algorithm performance. The second scenario could be more challenging and clarifies the real contribution of our algorithm. For example, the excitation values $P_1 = 0.045$, $P_2 = 0.044$, $N_1 = 0.0136$, $N_2 = 0.0138$, $N_3 = 0.4$, $B_1 = B_2 = 0.685$ has the output vector (1.0,0.68,0) which is a right state. Observe that activation values of positive regions P_1 and P_2 are quite similar, however, the overall evidence against decision d_2 suggests accepting d_1. It could be also possible that two dominant positive regions have the same activation value being difficult to take a decision. In such cases the map computes the decision by using negative and boundary regions.

The last scenario takes place when multiple boundary regions are noted at the same time, but no positive region is activated (e.g. $N_1 = 0.4$ and $B_2 = B_3 = 0.6$). This initial state leads to the output (0,0.54,0.54) where choices d_2 and d_3 are equally adequate, but it definitely suggest rejecting d_1. In such cases the decision-maker could adjust the similarity threshold with the goal of reducing the cardinality of boundary regions, and so reducing the overall uncertainty over the system. But if no change is observed then both decisions should be equally considered by experts. Following figure 2 illustrates the network behavior for all synthetic scenarios discussed above, where the activation degree of each decision neuron (over the time) is plotted.

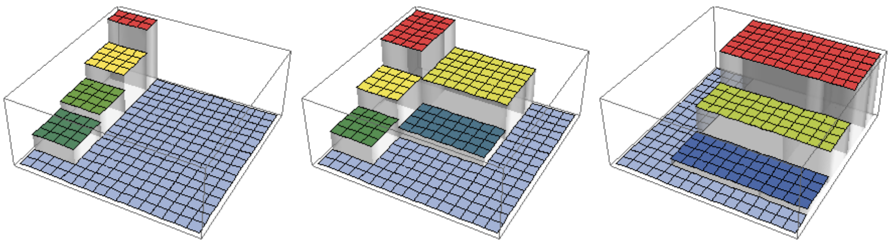

Fig. 2. Activation value of decision neurons for different scenarios. i) a single positive region is activated, (ii) two positive regions are activated, (iii) only boundaries regions are activated.

4.1 Decision-Making in Travel Behavior Problems

The transport management appears in all modern societies due to the cost that implies and because of the importance for social and economic process in a country. In such situations users intuitively select the most convenient transport mode mainly based on their expertise about a wide range of context variables such as: temperature, saving money, bus frequency, precipitation, physical comfort, etc.

Recently, M. León et al. [17] proposed a modeling based on Cognitive Mapping theory to characterize mental representations of users. Moreover, a learning algorithm based on Swarm Intelligence for tuning the system causality was developed, allowing the simulation of observed patterns. It facilitates the formulation of new strategies that efficiently manage existing resources according to the customer's preferences. Being more explicitly, each variable is evaluated by experts and then the best transport mode is selected (i.e. bus, car or bike). In order to evaluate the prediction capability of our model, we compare the accuracy on predicting the most adequate choice against other well-known algorithms such as: Multilayer Perceptron (MLP), Decision Trees (DT), Bayesian Networks (BN), and a Fuzzy Cognitive Map (FCM).

Before presenting experimental results a similarity relation to determine when two objects belong to the same similarity class is required. Following equation shows the symmetric relation R used in this work, where x and y are two objects of the universe of discourse, A is the set of attributes describing the task, m_i and M_i are the lower and upper values for the ith variable respectively, whereas the factor $0 < \varepsilon \leq 1$ represents the similarity threshold. Notice that, if $\varepsilon = 1$ then the relation R is reflexive, transitive and symmetric leading to the Pawlak's model for discrete problems. It is also possible to incorporate a weight for each attribute denoting the relevance degree of each variable which provides accurate approximations. However, in all simulations performed next, we assume that attributes variables have the same importance.

$$R: xRy \Leftrightarrow \frac{1}{|A|} \sum_{i=1}^{|A|} \left(1 - \frac{|x_i - y_i|}{M_i - m_i} \right) \geq \varepsilon \tag{4}$$

Table 1 shows the averaged prediction accuracy of RCN against four models taken from references [18]. It involves the accuracy for both optimistic and pessimistic trials over 220 knowledge bases concerning travel behavior problems. Toward this goal two studies were designed. In the first study (E1) stored scenarios serve either for training and validation (empirical error). The optimistic test is necessary because it reflects the self-consistency of the method, a prediction algorithm certainly cannot be deemed as a good one if its self-consistency is poor. In the second study (E2) testing cases were used to obtain a pessimist estimation (real error) using a cross- validation process with 10-folds. In such experiments "predicting the most important decision" means to find the transport mode having better expected utility, while "predicting the correct order in decisions" means to establish a proper ranking among decisions.

In case of the RCN model the similarity threshold ε is fixed to 0.9. This process is performed by "trial and error" although we could use a learning method as was stated in the previous section. However, in the present paper the authors prefer to be focused on the methodology to deal with decision-making problems.

Table 1. Prediction accuracy achieved by selected algorithms for both studies

	Predicting the most important decision					Predicting the correct order in decisions				
Study	MLP	BN	DT	FCM	RCN	MLP	BN	DT	FCM	RCN
E1	97.38	95.63	94.26	99.47	99.50	94.38	93.12	87.29	96.27	95.11
E2	92.06	91.37	89.39	93.74	93.29	82.40	80.25	77.59	88.72	90.32
Av	94.72	93.50	91.82	96.60	95.89	88.39	86.68	82.44	92.45	92.71

From Table 1 we can conclude that the RCN model is an appropriate alternative for addressing decision-making tasks having continuous (or mixed) features. It showed good prediction ability, outperforming traditional approaches such as MLP, BN or DT which have proved to be very competent classifiers. What is more, for this study case, our network performs comparable regarding the FCM model. It should remarked that the FCM topology introduced in [17] is problem-dependent, designed for addressing decision-making problems concerning public transportation issues, hence it cannot be generalized to other application domains. On the contrary, the RCN topology does not require specific information about the decision problem.

5 Conclusions

Decision-making problems have become an interesting research direction mainly due to their complexity and practical applications. In such problems experts express their preferences about multiple feasible alternatives according to the problem attributes or descriptors; the decision-making method must to derive a solution.

A relevant issue in decision-making tasks is related to the knowledge quality, since experts frequently have different perceptions for the same situation. To deal with this drawback some RST-based algorithms have been developed (e.g. three-way decision rules). This paper proposed a novel approach which combines the capability of Rough Sets for handling inconsistent patterns, with the simulation features of Sigmoid Fuzzy Cognitive Maps. It results in a new hybrid model, called Rough Cognitive Networks, which allows to solve difficult decision-making problems having discrete, continuous or mixed attributes. This model comprises three stages: (i) the estimation of positive, negative and boundary regions, (ii) the construction of the map topology, and (iii) the exploitation phase using the similarity class of the target instance.

During numerical simulations we observed that RCN are capable of computing the expected decision in scenarios where positive or boundary regions have similar (even identical) activation value. The main reason behind this positive result arises from the map inference process, which iteratively accentuates a pattern. The sigmoid threshold function also plays an important role in the simulation phase because it quantifies the activation degree of decision neurons. Similarly, using all computed regions allows to select the appropriate output from decisions having similar activation.

In a second study we compare the performance of our algorithm against traditional classifiers by using a real-life study case. From this study we noticed that RCM have high prediction capability, even without using specific knowledge about the decision problem. The future work will be focused on exploring the accuracy of such networks when facing more challenging classification problems.

References

1. Pawlak, Z.: Rough sets. Int. J. of Information and Computer Sciences 11, 341–356 (1982)
2. Bello, R., Verdegay, J.: Rough sets in the Soft Computing environment. Information Science 212, 1–14 (2012)
3. Pérez, R.B., Garcia, M.M.: Probabilistic approaches to the Rough Set Theory and their applications in decisionmaking. In: Espin, R., Pérez, R.B., Cobo, A., Marx, J., Valdés Olmos, R.A. (eds.) Soft Computing for Business Intelligence. SCI, vol. 537, pp. 67–80. Springer, Heidelberg (2014)
4. Yao, Y.: Three-way decisions with probabilistic rough sets. Information Science 180, 341–353 (2010)
5. Yao, Y.: Three-way decision: An interpretation of rules in rough set theory. In: Wen, P., Li, Y., Polkowski, L., Yao, Y., Tsumoto, S., Wang, G. (eds.) RSKT 2009. LNCS (LNAI), vol. 5589, pp. 642–649. Springer, Heidelberg (2009)
6. Wong, S., Ziarko, W.: Algorithm for inductive learning. Bulletin of the Polish Academy of Sciences Technical Sciences 34, 271–276 (1986)
7. Yao, Y.: The superiority of three-way decisions in probabilistic rough set models. Information Science 180, 1080–1096 (2011)
8. Kosko, B.: Fuzzy Cognitive Maps. Int. Journal of Man-Machine Studies 24, 65–75 (1986)
9. Kosko, B.: Hidden patterns in combined and adaptive knowledge networks. International Journal of Approximate Reasoning 2, 377–393 (1988)
10. Kosko, B.: Fuzzy Engineering. Prentice-Hall Inc., New York (1997)
11. Bueno, S., Salmeron, J.L.: Benchmarking main activation functions in Fuzzy cognitive maps. Expert Syst. Appl. 36, 5221–5229 (2009)
12. Tsadiras, A.K.: Comparing the inference capabilities of binary, trivalent and sigmoid fuzzy cognitive maps. Information Science 178, 3880–3894 (2008)
13. Slowinski, R., Vanderpooten, D.: A generalized definition of rough approximations based on similarity. IEEE Transactions on Data and Knowledge Engineering 12, 331–336 (2000)
14. Filiberto, Y., Bello, R., et al.: A method to build similarity relations into extended rough set theory. In: Proceedings of the 10th International Conference on Intelligent Systems Design and Applications, ISDA 2010, pp. 1314–1319. IEEE (2010)
15. Wilson, D.R., Martínez, T.R.: Improved heterogeneous distance functions. Journal of Artificial Intelligence Research 6, 1–34 (1997)
16. Miao, Y., Liu, Z.Q.: On causal inference in fuzzy cognitive maps. IEEE Transactions on Fuzzy Systems 8, 107–119 (2000)
17. León, M., Nápoles, G., García, M.M., Bello, R., Vanhoof, K.: Two Steps Individuals Travel Behavior Modeling through Fuzzy Cognitive Maps Pre-definition and Learning. In: Batyrshin, I., Sidorov, G. (eds.) MICAI 2011, Part II. LNCS (LNAI), vol. 7095, pp. 82–94. Springer, Heidelberg (2011)
18. León, M., Nápoles, G., Bello, R., Mkrtchyan, L., Depaire, B., Vanhoof, K.: Tackling Travel Behaviour: An approach based on Fuzzy Cognitive Maps. International Journal of Computational Intelligence Systems 6, 1012–1039 (2013)