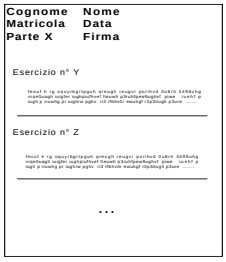
Soft Computing 3-09-2010

- Use only the answer sheets provided to you, write clearly on the top left hand corner (**like in the figure**) firstly the surname, then the name (in BLOCK characters), on the following line enrolment number and date, on the following line part of the exam (Part 1 or Part 2) and signature
- Hand in the two parts in separate sheets
- Indicate the number of each exercise and separate each solution from each other clearly with a horizontal line
- Write CLEARLY by pen or pencil



COMMENT: some people still did not follow the above reported instructions. From this exam, 2 points depend on the accomplishment of what mentioned above. In particular: all the information have to be written on the left part of each sheet, first the surname than the name, surname and name have to be written clearly (in particular they would usually be different from how you sign), all the requested data have to be included in the given order. Please, remember also that your goal in writing your solution is to transmit information to your examiner, not just to fill the sheet in the shortest time, so, the more your writing is clear, the higher the probability that who is evaluating you can understand your message.

Part 1

1.1. Fuzzy Design [7/32]

Design a fuzzy system able to control some aspects of the washing cycle of a washing machine. The system considers the intervals between two activations of the electrovalve loading water, related to the degree of water absorption of clothes. Longer time intervals correspond to less absorbing materials, which are considered as more . The system modifies the time of cloth handling, as well as the centrifuge speed, according to the care need of clothes.

Select the input and the output variables of the fuzzy system, define the corresponding fuzzy sets, **justifying** shape and position, and at least 3 rules for modeling the problem.

Solution

Many solutions are possible. We report here only one, with remarks about correct and incorrect choices.

There is just one input variable, which might be either the average interval between two input valve activations or a qualification of the material, which was assumed in the text to be derived from the first one. No other information about the type of clothes is available.

The output is given by two separate variables, that control two different aspects of the washing cycle: the amount of movement during washing and rinsing, and the speed of the centrifuge.

The shape of outputs can be singletons if the defuzzyfication method is center-of-mass, so that we can exploit all the values and we have a proportional control. With triangles or trapezoids we should select a different defuzzyfication method and also justify why should we have them (writing that so we favor rules with low fitness is not enough: why should we like to have this?) The number of output MF have to be consistent with the number of combinations of the input values. It makes no sense to have 3 possibilities in input and 7 values for each variable (which are independent) in output.

1.2. Genetic algorithms [2/32]

Explain how is typically represented a model in Genetic Programming, and how genetic operators are implemented in this case.

Solution

This is just a theory question whose answer can be found in books and slides. Genetic Programming is a special case of Genetic Algorithm, where the solution is coded as a tree connecting operands and language words such as "+", "IF", "WHILE", etc. The crossover in this case is done by selecting a cutting point in the parents and exchanging subtrees. Mutation can either mutate the value of the operands or the operators themselves, eventually (it depends from the language) checking the consistence of the syntax obtained.)

Many people described the generic Genetic Algorithm model, which is not what was asked in the question.

1.3. Reinforcement learning systems [7/32]

Design a Reinforcement learning system to learn the optimal policy of a washing machine to optimize energy and water use, by maintaining an acceptable washing quality. let us assume that the situation the washing machine can be in are fixed a priori, that they be in a finite number, and that they be defined by the state of the machine, (ectrovalve controlling the water input open/close, slow handling (for washing and rinsing)), fast movement (for centrifuge)) and time spent in a given state (discretized with 1 minute granularity). Possible actions are those allowing to pass from one state to another, and are selected every minute; among them we have also the one doing nothing (NOP). Finally, measures of water and energy use are available, as well as a 3-valued evaluation of the washing quality (good, acceptable, unacceptable)

Model the problem, specifying the **states**, the possible **actions** (**that enable the agent to pass from one state to another**), and a **reinforcement function**. Justify the choice of a reinforcement distribution algorithm for this application,

on the basis of the selected model.

Solution

Many solutions are possible, He we report one of them.

The states are described by the following variables: $input_valve_state \in \{Open, Close\}$, $movement \in \{washing/rinsing, centrifuge\}$, $time \in \{0..100\}$

The actions bring from one state to the other: NOP (increases time with the same values of the other variables), OPEN/CLOSE input valve, START-SLOW-MOVEMENT, STOP-MOVEMENT, START-FAST-MOVEMENT.

The reinforcement function gives positive reward when the quality of washing is acceptable, with a contribution inversely proportional to the amount of energy and water used.

The reinforcement distribution algorithm can be $Q(\lambda)$ *with eligibility traces, considering that reinforcement is given only at the end of the washing cycle, where it is possible to evaluate the elements of the reinforcement function. In this case the model is well known (each action changes deterministically the state), and, in principle, Dynamic Programming could be used, although reinforcement is given only in one state and information can be not enough to converge. Motecarlo algorithms have few sense due to sparse reinforcement and the real-world application.*

Part 2

2.1 Neural Networks [Score 7/32]

Lets consider the neural network model with feedforward topology having I input, J hidden neurons and 1 output. To reduce outliers influence during training it is preferable to use a difference noise model instead of the Gaussian one. A possible choice is the Cauchy distribution:

$$p(x|\mu,\gamma) = \frac{1}{\pi\gamma[1+(\frac{x-\mu}{\gamma})^2]} = \frac{1}{\pi}\left[\frac{\gamma}{(x-\mu)^2+\gamma^2}\right].$$

- derive the error function corresponding to the maximum likelihood estimation hen using the Cauchy error model
- derive the backpropagation formulas for the network weights in the case of Cauchy error model
- what is overfitting? Does this kind of model soffer it? If so, describe the two techniques presented during the course to deal with it.
- what kind of activation function should we use for the neurons (input, hidden, and output) of this network when facing a regression problem? What about classification? Explain your answers.

2.2 Bayesian Classifier [Score 4/32]

A recent survey has studied the effects of pre and extramarital sexual activity on divorce. The observed variables are: Gender, PreSex, ExtraSex, and the output class is Marital status. These data are reported in in the table with rows matching the same configuration aggregated in the count column.

GENDER	PRE	EXTRA	MARITAL	COUNT
Men	No	No	Divorced	68
Women	No	No	Divorced	214
Men	Yes	No	Divorced	60
Women	Yes	No	Divorced	54
Men	No	Yes	Divorced	17
Women	No	Yes	Divorced	36
Men	Yes	Yes	Divorced	28
Women	Yes	Yes	Divorced	17
Men	No	No	Married	130
Women	No	No	Married	322
Men	Yes	No	Married	42
Women	Yes	No	Married	25
Men	No	Yes	Married	4
Women	No	Yes	Married	4
Men	Yes	Yes	Married	11
Women	Yes	Yes	Married	4

- describe the maxima a posteriori classification approach (a.k.a. Bayesian Classifier) and derive from it the Naive Bayes Classifier
- write the complete Naive Bayes Classifier learnt from the data in the table (use Laplace estimator if needed)
- classify the record [???,Yes,Yes]

2.3 Bayesian Networks [Score 4/32]

"You have a new burglar **alarm** installed in your house. It is fairly accurate at detecting a **burglary**, but also sounds occasionally when there is an **earthquake** (a vary rare event, but still possible). Two neighbors, **John** and **Mary** have promised to call, when they hear an **alarm**. **John** sometimes confuses your telephone ring for an **alarm** and **Mary** doesn't always hear the **alarm** because of loud music playing.

Considering bold events as the variables of your problem, answer the following questions:

- draw the Bayesian network corresponding to the situation previously described including (your) conditional probability tables
- describe all independent and conditionally independent variables
- compute the (exact) probability of an alarm given that both John and Mary call you
- is it possible to use Message Passing to compute the previous probability? What about Monte Carlo Sampling? And What about Likelihood Sampling? Which one out of these three would you chose and why?