ATCI Final Exam
Date: 9/6/2017
Duration: 12h to 14h

Answer the exam in four different pages: questions from 1 to 5, 6 to 8, 9 to 10, and 11 to 15. Do not forget to write your name in ALL the sheets.

1. What differentiates the training processes of FIR and NN approaches?

2. How can you deal with the curse of dimensionality problem when working with FIR?

3. How does FIR solve the uniqueness problem in the discretization process? What does it differ with respect the traditional fuzzy approximation?

4. Which are the main advantages and disadvantages of Pittsburgh and Michigan approaches?

5. Describe briefly the iterative rule learning algorithm. Give an example of useful fitness function for this approach.

6. Let us consider learning a task that involves static input/output data, where the different examples are independent pairs of input and target vectors. Is there any advantage of applying a recurrent neural network trained by recurrent backpropagation in front of applying a multi-layer perceptron trained by standard backpropagation? If yes, which are they? If not, which are the disadvantages?

7. Do the Time-Delay Neural Networks suffer from the vanishing gradient problem? What feature of these networks limits their capacity of taking into account relevant information in the past input sequence?

8. In the Long-Short Term Memory (LSTM) architecture, what do you think is the practical difference between the following two network design options:

   a) a hidden layer consisting of a single memory block containing 3 memory cells;

   b) a hidden layer consisting of 3 memory blocks containing a single memory cell each?

   Justify your answer based on the roles of the input, forget and output gates.

9. What is the representation space in a Similarity Neural Network (SNN)? What advantages do you think it can bring over operating on the input space? Can you name a disadvantage?

10. How do SNNs (based on Gower's similarity measure) deal with missing values? Can this process be seen as an imputation? If so, what would be the imputed value? If not, why not?

11. In constructive neural networks, the selection of the input-layer weights is a hard non-linear problem and cannot be solved analytically. Enumerate three different ways to obtain them in practice.

12. Both Deep Networks and Support Vector Machines can be seen as models that work with feature representations of the data. What is the main difference between the feature representations of both models?

13. What are the main ideas of greedy layer-wise unsupervised pre-training? Which learning models do you know that can be used to implement it?

14. What is the most important consequence that P(h|x) and P(x|h) factorize in Restricted Boltzmann Machines? Does it have any practical application?

15. What is the main motivation for using the convolution operation?