A Neuro-Fuzzy approach to identify a Hierarchical Fuzzy System for modelling Aviation Pilot Attention

DI NUOVO, Alessandro, CANNAVO, Rosario, DI NUOVO, Santo, TRECCA, Giorgia and RAVECCA, Fabio

Published version

DI NUOVO, Alessandro, CANNAVO, Rosario, DI NUOVO, Santo, TRECCA, Giorgia and RAVECCA, Fabio (2016). A Neuro-Fuzzy approach to identify a Hierarchical Fuzzy System for modelling Aviation Pilot Attention. In: IEEE World Congress on Computational Intelligence 2016, IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2016), Vancouver, Canada, 24-29 July.

A Neuro-Fuzzy approach to identify a Hierarchical Fuzzy System for modelling Aviation Pilot Attention

Alessandro Di Nuovo
Department of Computing
Sheffield Hallam University, Sheffield, UK
Email: a.dinuovo@shu.ac.uk

Rosario Cannavó
Cognitive Technologies and Services Srl
Catania, Italy
Email: rosario.cannavo@cogtech.it

Santo Di Nuovo
Department of Education Sciences
University of Catania, Catania, Italy
Email:s.dinuovo@unict.it

Giorgia Trecca, Fabio Ravecca
Italian Navy, Rome, Italy
Email: [giorgia.trecca, fabio.ravecca]@marina.difesa.it

***Abstract*—Attention has been shown to be a predictor of flight performance and, therefore, it is necessary to assess this cognitive ability to evaluate candidate aviation pilots and to verify if the pilot has the sufficient attention level required for the flight duties.**

In this paper, we present a study that uses a Neuro-fuzzy approach to identify a benchmark model of the aviation pilot attention level. The model is learned from the data examples collected using a computerized battery of seven tests, which was specifically built and validated to assess the main cognitive factors related with the aviation pilot attention in realistic scenarios. Data examples were collected in experimental session with a total of 180 participants, 96 military aviation pilots and 84 untrained people as controls.

The aim is to build the model as a classifier that is able to discriminate between the two groups, allowing to identify the peculiar profile of the aviation pilot as opposed to control subjects.

Classification performance analysis shows that a hierarchical fuzzy system has a better accuracy than single stage classification algorithms and gives more details about the different attention factors. Moreover, a fuzzy system for our model because it can be readable by human instructors and used as guidance for the training.

The ultimate objective of our work is an expert system that will be able to assess the attention performance and compare it against the typical profile of the aviation pilot. This will be used as an instrument for more accurate selection and training of aviation pilots by identifying the areas of deficit that need to be improved, and to measure if a pilot has a sufficient level of attention before the flight.

I. INTRODUCTION

Attention is defined as *a concentration of mental activity that allows you to take in a limited portion of the vast stream of information available from both your sensory world and your memory* [1]. The measurement of the attention as a predictor of flight performance was previously shown in several works [2], [3] and theoretical and empirical evidence exists in support of the argument that the control of attention and the ability to establish better attention-management can develop with training [4].

Starting from this background, our work aims to identify a specific attention profile by measuring the peculiar attentional style of aviation pilots that should be acquired during the training and maintained during the duty. Here the attention profile is defined as the set of features that characterize, in our case, the pilots of aircraft and it is a stable trait of the person. The identification of that style and, moreover, of an instrument to objectively measure it, can give a two-fold contribution: (i) as a predictive tool for the selection of most promising candidates and to identify the factors that they should improve during the training; (ii) to develop an innovative procedure to verify, before and during the flight duties, that the pilot has the sufficient attention level to be qualified for that flight.

In this paper, we present a study that uses a Neuro-fuzzy approach to identify a benchmark model of the aviation pilot attention level. The model has been developed from the results of the administration of a battery of seven computer tests which has been demonstrated to be capable of discriminating between aviation pilots and untrained controls [5]. The battery was administered to a group of aviation pilots and a control group of untrained people. Data examples were collected in experimental session with a total of 180 participants, 96 military aviation pilots and 84 untrained people as controls. The aim is to build the model as a classifier that is able to discriminate between the two groups, therefore, to identify the peculiar profile of the aviation pilot as opposed to control subjects.

Details on the participants to our experimental sessions and on a brief description of the battery of tests are given in the Section II. Then, Section III gives a brief introduction of the computational intelligence and machine learning algorithm used to learn classification models that we used for benchmarking the accuracy of the fuzzy models. Section IV presents a preliminary performance analysis of all the algorithms presented in the previous section with numerical results and statistical analysis, and then the hierarchical fuzzy model for assessing the aviation pilot attention. Finally, Section V presents our conclusion and the direction for future work.

II. Material and Methods

The first subsection introduces the computerized test battery that has been previously validated for measuring attention of aviation pilots [5]. The second subsection details the experimental procedure we adopted to collect the dataset used in this work. The third subsection gives descriptive statistics of the participants.

A. Aviation pilot attention test battery

In this work, we used a battery of seven computerized attentive tests, which were specifically designed and validated for assessing the aviation pilot attention. This battery has been implemented in a single computer software solution, which simplifies the administration of the different tests and data collection. The battery has the characteristic to evaluate some peculiar factors that are crucial for aviation pilots. In particular, it differentiates among central, mid-peripheral and far peripheral focus of the visual attention. Peripheral vision is used to detect objects that are located at an eccentricity of more than nine degrees with respect to the foveal vision. "Far peripheral" vision exists at the edges of the field of view; "mid-peripheral" vision exists in the middle of the field of view. Test scenarios were designed using real-world pictures and videos to increase the level of ecological validity with real-life experience. Details about the tests battery and the software are given in [5]. Table I summarize the measures for each test.

B. Experimental setting and Procedure

The experimental study was conducted in computer rooms, where the attention test battery was administered to several subjects at same time. The pilots were selected among the staff of three bases of Italian Naval Aviation (MARISTAELI Catania, Luni, Grottaglie). The controls were selected among students and staff of the University of Enna Kore. Subjects were in good health, rested and comfortably seated in front of a computer monitor at a distance equal to their arm, this was to ensure that the lateral area of the monitor coincide with the peripheral area of the vision. Finally, in order to avoid auditory interference all subjects were wearing a headset during the whole duration of the test. This is because the presence of interfering stimuli may be partially distracted from the task, distorting the attentional profile the subject. Before each session a collective briefing was done to explain the general issues of the test battery. At the end of each session an individual debriefing was done to collect feedbacks from the subjects.

C. Participants

The number of subjects that completed all the tests in the battery is N=180, divided into two groups: 96 navy aviation pilots and 84 untrained healthy people as control group. Subjects' age was in the range of 18-53 years old. Pilots were selected among subjects that have completed their training and experienced at least 100 hours of flight (real or simulated). Controls were recruited, after a preliminary selection from a wider sample, excluding those who underwent particular trainings (e.g. sports or professional car driving experience), but with at least an education history that allows them to be enrolled in a pilot training course, i.e. high school degree. Because of these criteria a different mean age was found between the two groups, where controls are significantly younger (pilots' mean=30.44, std = 4.29; controls' mean=23.04, std=7.71, $p < 0.01$). The participants gave written informed consent to use the data collected during the trial for research purposes.

III. The Computational Intelligence and Machine Learning approaches used in this Work for Building the Classification Model

In the following we briefly presents the algorithms used in our experiments and their parameters. If not otherwise specified the learning algorithm and its parameters were the default implementation of MATLAB 2014b.

A. Artificial Neural Networks (ANNs)

. In this work, we considered two versions of Artificial Neural Networks.

The first is a pattern recognition neural network (PRNN), which is a three-layer feedforward networks that can be trained to classify inputs according to target classes. The capability of multi-layer feed-forward ANNs in creating models for arbitrary non-random input-output mappings has been firstly demonstrated in [6], [7]. Our feedforward multi-layer ANNs for pattern recognition comprises 10 neuronal units in the hidden layer. The hidden layer has the hyperbolic tangent sigmoid as the transfer function $\mathtt{tansig}(x) = 2/(1+\exp(-2*x))-1$, while the output layer has competitive neurons. In a competitive layer the output will be 1 for the most active neuronal unit and 0 for all the others. We used a gradient descent with momentum weight and bias as the learning function, and Levenberg-Marquardt back propagation as the training function [8], [9]. The number of neurons in the hidden layer was chosen as 10, which maximise the accuracy and reliability of the model being considered.

The second is a learning vector quantization neural networks (LVQNN), which consist of two layers. The first layer maps input vectors into clusters that are found by the network during training. The second layer merges groups of first layer clusters into the classes defined by the target data. The network is trained applying a winner-take-all Hebbian learning-based approach [10]. An LVQ model comprises several vector prototypes which are defined in the feature space of observed data. The training algorithm determines, the prototype which is closest to the input according to a given distance measure. The position of the winner prototype is then updated, i.e. it will move closer if it correctly classifies the data point or moved away if it classifies the data point incorrectly.

B. Support Vector Machines (SVM)

SVM are non-probabilistic binary linear classifiers. An SVM classifies data by finding the best hyperplane that separates all data points of one class from those of the other

TABLE I
SUMMARY OF THE DATA ANALYSED. ALL VARIABLE SCORES ARE NEGATIVES, I.E. THE LOWER THE BETTER.

Test and variable name	Test Short ID	No. of Stimuli	Omissions	Errors	Median Reaction time (MEDTIME)
SIMPLE REACTION TIME					
CENTRAL	1A	40	✓	✓	✓
PERIPHERAL	1B	40	✓	✓	✓
MULTIPLE SEARCH	2	20	✓	✓	✓ Total TIME
COLOR-WORD INTERFERENCE					
COLOR DISCRIMINATION	3A	40	✓	✓	✓
COLOR-WORD	3B	40	✓	✓	✓
GROUND INTERFERENCE	4	16	✓	✓	✓
DIVIDED ATTENTION					
AUDITORY	5A	16	✓	✓	✓
VISION	5V	22	✓	✓	✓
DIGIT SPAN					
DIRECT	6D	8	✓	-	-
INVERSE	6I	8	✓	-	-
GLOBAL VISION	7	18	✓	✓	✓
CENTRAL	7C	6	✓	-	✓
MID-PERIPHERAL	7M	6	✓	-	✓
FAR-PERIPHERAL	7F	6	✓	-	✓

class. The best hyperplane for an SVM means the one with the largest margin between the two classes. Margin means the maximal width of the slab parallel to the hyperplane that has no interior data points. Support Vector Machines can be used when exactly two classes can be identified in the data. The SVM is trained to classify vectors x according to the following equation:

$$c = \sum_i \alpha_i k(s_i, x) + b$$

where s_i are the support vectors, α_i are the weights, b is the bias, and k is a kernel function. In the case of a linear kernel, k is the dot product. If $c \geq 0$, then x is classified as a member of the first group, otherwise it is classified as a member of the second group. More details can be found in [11].

C. Classification trees (CTREE)

Classification or Decision trees predict responses to data [12]. To predict a response, follow the decisions in the tree from the root (beginning) node down to a leaf node. A classification tree is a flow-chart-like structure, where each internal (non-leaf) node denotes a test on an attribute, each branch represents the outcome of a test, and each leaf (or terminal) node holds a class label. The topmost node in a tree is the root node. The leaf node contains the response. Classification trees give responses that are nominal, such as 'true' or 'false'. To learn the classification tree we used the standard CART algorithm [13], which is one of the main algorithms for constructing classification trees from class-labelled training tuples. The algorithm creates a multi-way tree, finding for each node (i.e. in a greedy manner) the categorical feature that will yield the largest information gain for categorical targets. Pruning is done by removing a rules precondition if the accuracy of the rule improves without it. CART constructs binary trees using the feature and threshold that yield the largest information gain at each node.

D. Adaptive-Network-based Fuzzy Inference Systems (ANFIS)

ANFIS are a class of adaptive networks that are functionally equivalent to fuzzy inference systems (FIS) originally proposed in [14]. Using a given input/output data set ANFIS constructs a FIS whose membership function parameters are tuned (adjusted) using a backpropagation algorithm alone in combination with a least squares type of method. This adjustment allows your fuzzy systems to learn from the data they are modelling. In detail for premise parameters that define membership functions, ANFIS employs gradient descent to fine-tune them. For consequent parameters that define the coefficients of each output equations, ANFIS uses the least-squares method to identify them. This approach is thus called hybrid learning method since it combines gradient descent and the least-squares method. Before the ANFIS training phase, it must be specified an initial FIS model structure of the Takagi-Sugeno type [15]. To specify the initial model structure there are various approaches, in this work we considered two clustering approaches that are best suited for high dimensional databases. They are:

- *Subtractive Clustering* (ANFIS-SUB), that generates the initial FIS model for ANFIS training by first applying subtractive clustering on the data [16].
- *Fuzzy C-Means Clustering* (ANFIS-FCM), that generates the initial FIS using fuzzy c-means (FCM) clustering by extracting a set of rules that models the data behaviour. Details on FCM are in [17].

The hierarchical fuzzy model was build using a two step procedure: first seven fuzzy rule based classifier systems were

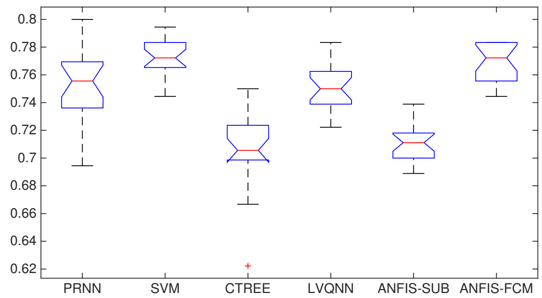

Fig. 1. Classification comparison: Boxplot of Correct Classification rates over 21 random repetitions of 10-fold cross-validation. ANFIS-FCM and SVM approaches were able to build the best classifiers.

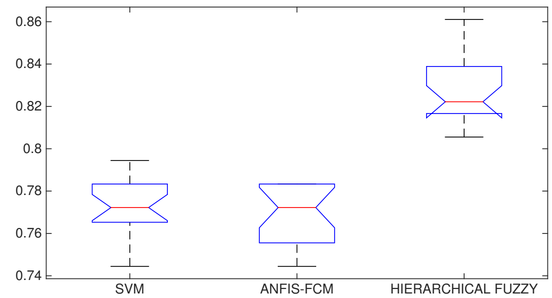

Fig. 2. Classification comparison Hierarchical Model: Boxplot of Correct Classification rates over 21 random repetitions of 10-fold cross-validation. The hierarchical architecture clearly outperforms the standard algorithms.

learned using ANFIS, one for each sub-test of the battery; then the outputs of all classifiers become the inputs to train the second layer of the hierarchy, while the targets were the actual classes.

IV. Results and Discussion

The modelling methods presented above were applied to the dataset of examples collected from the experimental session with the 180 participants as detailed in the Section II. Numerical results were obtained by applying 21 times the *ten-fold cross validation* test, in which each model was executed ten times, first applying the algorithm to a random subset of 162 examples for learning and then validated on the remaining 18 elements. The values presented are those calculated during the validation.

TABLE II
Classification Comparison on the entire dataset

Algorithm	Correct Classification Rate			Precision (Medians)	
	Min	Max	Median	Pilots	Controls
PRNN	*69.44%*	*80.00%*	*75.56%*	*80.21%*	*71.08%*
SVM	*74.44%*	*79.44%*	*77.22%*	*86.46%*	*66.67%*
CTREE	*62.22%*	*75.00%*	*70.56%*	*70.83%*	*70.24%*
LVQNN	*72.22%*	*78.33%*	*75.00%*	*86.46%*	*59.52%*
ANFIS-SUB*	*68.89%*	*73.89%*	*71.18%*	*89.58%*	*50.00%*
ANFIS-FCM*	*74.44%*	*78.33%*	*77.22%*	*82.29%*	*71.43%*

*Note that, given the high dimensionality of the dataset (43 features), it was not possible to apply the ANFIS optimisation as there were not enough examples. The results shown in this table refer to the fuzzy system built with the cluster algorithm only.

Table II presents the main statistical descriptive variables - minimum (min), maximum (max), median - for each algorithm considered in our experiments. Best results are achieved by ANFIS-FCM and SVM, while the others are significantly inferior. Note that it was not possible to apply the ANFIS algorithm to the entire database because of the high dimensionality with respect to the relatively low number of examples. For this preliminary study, the results are from the application of the clustering algorithms only.

According to the Kruskal-Wallis test p is < 0.001, thus there is a statistically significant difference between the algorithm results. The Kruskal-Wallis test ranks SVM first (97.45) and ANFIS-FCM second (94.24), but there is no statistical difference between these two result distributions ($p = 0.997$). The two algorithms have a statistically significant difference with all the others. In other words, the best performance in terms of Correct Classification rate is achieved by both SVM and ANFIS-FCM. However, if we consider the classification accuracy of the single groups, we see that ANFIS-FCM is significantly more accurate with the controls, thus we consider it more reliable to build a model for discriminating the two groups. Figure 1 reports the box plot for the six distributions obtained running 21 times the 10-fold cross validation for each algorithm with different seeds.

The better performance of the ANFIS approach was expected, indeed, some of the authors obtained similar results in a previous study in which ANFIS was employed for function approximation [18].

Table III reports the results for ANFIS-FCM applied to single tests. Good results can be seen in the identification of aviation pilots, especially for tests n.1 (Simple reaction time) and n.7 (Global Vision). These results confirms that the reaction time is a critical factor for measuring the attention and, therefore, a strong performance is a peculiar for aviation pilots. Other test are less accurate, e.g. Ground Interference, this is in-line with the previous results of [5].

Table III includes the result of the Hierarchical model, which includes a second layer of fuzzy rules that receives as input the classification (output) of the single tests fuzzy systems, then processes and combine the results to improve the final classification. Indeed, our results show an increased accuracy, over 80% for both groups. Comparing the Hierarchical model distribution with the other models, its better result is statistically significant with $p < 0.01$ obtained with the Kruskal-Wallis test. The box plot for SVM, ANFIS-FCM and the Hierarchical fuzzy model is depicted in figure 2.

The Hierarchical fuzzy model is schematically represented in figure 3. From the results we can state that the procedure to split the computation between two layers of a hierarchical fuzzy system significantly improves the classification accuracy of the system.

TABLE III
SINGLE TEST AND HIERARCHICAL SYSTEM CLASSIFICATION RESULTS WITH ANFIS-FCM, 10 FOLD CROSS-VALIDATION

Test	Fuzzy Rules N.	Classification rate			Precision (Medians)	
		Min	Max	Median	Pilots	Controls
1 SIMPLE REACTION TIME	3	*66.11%*	*71.11%*	*67.78%*	*79.17%*	*54.76%*
2 MULTIPLE SEARCH	2	*56.67%*	*63.33%*	*60.56%*	*71.88%*	*47.62%*
3 COLOR WORD INTERFERENCE	4	*62.78%*	*67.78%*	*66.11%*	*77.08%*	*53.57%*
4 GROUND INTERFERENCE	3	*48.33%*	*56.11%*	*53.89%*	*64.58%*	*41.67%*
5 DIVIDED ATTENTION	3	*55.00%*	*62.22%*	*58.89%*	*71.88%*	*44.05%*
6 DIGIT SPAN	13	*58.33%*	*66.11%*	*62.78%*	*69.79%*	*54.76%*
7 GLOBAL VISION	4	*66.67%*	*74.44%*	*71.11%*	*86.46%*	*53.57%*
HIERARCHICAL MODEL	2*	*77.22%*	*0.8278%*	*82.22%*	*83.33%*	*80.95%*

*This is the number of rules in the second layer of the hierarchy. See Figure 3 for details.

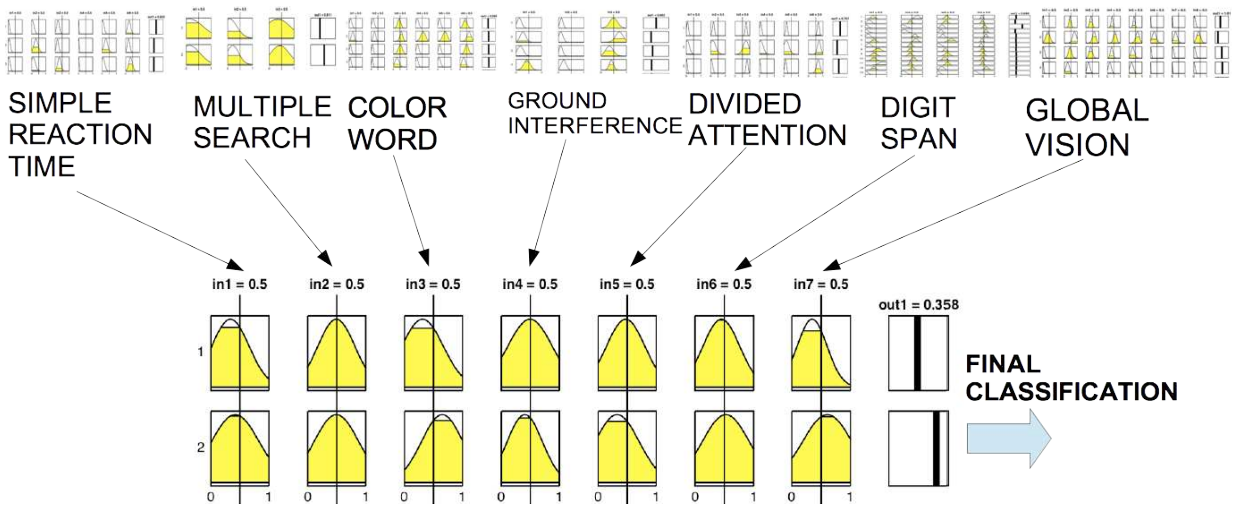

Fig. 3. A schematic representation of the Hierarchical Fuzzy Model. The first layer of the hierarchy is a collection of seven fuzzy rule based classifier systems, one for each subtest of the battery. The second layer receives the outputs (fuzzy) from the first layer and then computes the final class. This procedure significantly increases the performance of the classifier system.

V. CONCLUSION

This paper presents a study to identify the peculiar attentional model of aviation pilots. To this end, a battery of computerized tests was administered to a significant (N = 180) sample of pilots (N = 96) and untrained people (N = 84). Data collected were used as examples for supervised learning of computational models by means of machine learning and computational intelligence techniques. From the numerical experiments and comparison analysis we see two main results: (i) The fuzzy model is equally or more accurate and reliable than other obtained with neural network and machine learning approaches; (ii) The hierarchical learning structure significantly improves the classification accuracy, with the additional benefit that it gives specific details on single tests, which make possible to calculate the intermediate results and identify areas of strength and weakness.

In practice, the hierarchical fuzzy model can be used to assess the likelihood of candidate pilots by predicting their performance, and to increase the efficiency of their training by suggesting which factor they should improve. Indeed, the sub-models for each test can be used by instructors to evaluate the different attention factors and focus the training on those that represent a weakness.

Further development will focus on the definition of a battery of tests that can be integrated with onboard instrumentation (e.g. with fuzzy based augmented cognition systems like the one proposed in [19], [20]). This safety enhancer technology could evaluate real-time the current attention level before the start of a flight duty, and to assure that a sufficient level is always attained to fulfil with the highest safety requirements.

ACKNOWLEDGMENT

Authors want to thank CV A. Managó, CF M. Speranza and TV N. Casarini of the Italian Navy for their support to the administration of the computerized software to navy pilots. The research was partially supported by the Cognitive Technologies and Services SRL (also known as CogTech) with the implementation of the computer software (Copyright registered, 2013).

REFERENCES

[1] M. W. Matlin, *Cognition*, 8th ed. Wiley, 2012.
[2] R. North and D. Gopher, “Measures of attention as predictors of flight performance,” *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 18, no. 1, pp. 1–13, 1976.

[3] D. Gopher, “A Selective Attention Test as a Predictor of Success in Flight Training,” *Human Factors : The Journal of the Human Factors and Ergonomics Society*, vol. 24, no. 2, pp. 173–183, 1982.
[4] D. Gopher, M. Well, and T. Bareket, “Transfer of skill from a computer game trainer to flight,” *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 36, pp. 387–405, 1994.
[5] R. Cannavò, D. Conti, and A. D. Nuovo, “Computer-aided assessment of aviation pilots attention: Design of an integrated test and its empirical validation,” *Applied Computing and Informatics*, vol. 12, no. 1, pp. 16 – 26, 2016.
[6] G. Cybenko, “Approximation by superpositions of a sigmoidal function,” *Mathematics of Control, Signals and Systems*, vol. 2, no. 4, pp. 303–314, 1989.
[7] K. Hornik, M. Stinchcombe, and H. White, “Multilayer feedforward networks are universal approximators,” *Neural Networks*, vol. 2, no. 5, pp. 359–366, 1989.
[8] M. Hagan and M. Menhaj, “Training feedforward networks with the marquardt algorithm,” *IEEE Transactions on Neural Networks*, vol. 5, no. 6, pp. 989–993, Nov 1994.
[9] S. He, N. Sepehri, and R. Unbehauen, “Modifying weights layer-by-layer with levenberg-marquardt backpropagation algorithm,” *Journal of Intelligent Automation and Soft Computing*, vol. 7, pp. 233–247, 2001.
[10] T. Kohonen, *Learning vector quantization*. Springer, 1995.
[11] N. Cristianini and J. Shawe-Taylor, *An introduction to support vector machines and other kernel-based learning methods*. Cambridge university press, 2000.
[12] J. R. Quinlan, “Induction of decision trees,” *Machine learning*, vol. 1, no. 1, pp. 81–106, 1986.
[13] L. Breiman, J. Friedman, C. J. Stone, and R. A. Olshen, *Classification and regression trees*. CRC press, 1984.
[14] J.-S. R. Jang, “ANFIS: Adaptive-network-based fuzzy inference systems,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 23, no. 3, pp. 665–685, May 1993.
[15] T. Takagi and M. Sugeno, “Fuzzy identification of systems and its application to modeling and control,” *IEEE Transactions on System, Man and Cybernetics*, vol. 15, pp. 116–132, 1985.
[16] S. Chiu, “Fuzzy model identification based on cluster estimation,” *Journal of Intelligent & Fuzzy Systems*, vol. 2, no. 3, pp. 267–278, 1994.
[17] J. C. Bezdek, *Pattern Recognition with fuzzy objective function algorithms*. New York: Plenum Press, 1981.
[18] A. Di Nuovo, S. Di Nuovo, and S. Buono, “Intelligent quotient estimation of mental retarded people from different psychometric instruments using artificial neural networks,” *Artificial Intelligence in Medicine*, vol. 54, no. 2, pp. 135–145, 2012.
[19] A. G. Di Nuovo, R. B. Cannavo, and S. Di Nuovo, “An agent-based infrastructure for monitoring aviation pilot’s situation awareness,” in *IEEE Symposium on Intelligent Agents (IA)*. IEEE, apr 2011, pp. 1–7.
[20] A. G. Di Nuovo, R. B. Cannavò, and S. Di Nuovo, “An intelligent infrastructure for in-flight situation awareness of aviation pilots,” in *Foundations of Augmented Cognition. Directing the Future of Adaptive Systems*. Springer Berlin Heidelberg, 2011, pp. 598–607.