



HAL
open science

How to solve a classification problem using a cooperative tiling Multi-Agent System?

Thibault Fourez, Nicolas Verstaevel, Frédéric Migeon, Frédéric Schettini,
Frédéric Amblard

► To cite this version:

Thibault Fourez, Nicolas Verstaevel, Frédéric Migeon, Frédéric Schettini, Frédéric Amblard. How to solve a classification problem using a cooperative tiling Multi-Agent System?. 20th International Conference on Practical Applications of Agents and Multi-Agent Systems (PAAMS 2022), Jul 2022, L'Aquila, Italy. pp.166-178, 10.1007/978-3-031-18192-4_14. hal-03776374

HAL Id: hal-03776374

<https://ut3-toulouseinp.hal.science/hal-03776374>

Submitted on 14 Sep 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

How to solve a classification problem using a cooperative tiling Multi-Agent System?

Thibault Fourez^{1,2}[0000-0003-1677-6623], Nicolas Verstaevl¹[0000-0002-7879-6681], Frédéric Migeon¹, Frédéric Schettini², and Frédéric Amblard¹

¹ Institut de Recherche en Informatique de Toulouse, Université de Toulouse, CNRS, Toulouse INP, UT3, UT1, Toulouse, France
<https://www.irit.fr>

² Citec Ingénieurs Conseil, Geneva, Switzerland

Abstract. Adaptive Multi-Agent Systems (AMAS) transform dynamic problems into problems of local cooperation between agents. We present *smapy*, an ensemble based AMAS implementation for mobility prediction, whose agents are provided with machine learning models in addition to their cooperation rules. With a detailed methodology, we propose a framework to transform a classification problem into a cooperative tiling of the input variable space. We show that it is possible to use linear classifiers for online non-linear classification on three benchmark toy problems chosen for their different levels of linear separability, if they are integrated in a cooperative Multi-Agent structure. The results obtained show a significant improvement of the performance of linear classifiers in non-linear contexts in terms of classification accuracy and decision boundaries, thanks to the cooperative approach.

Keywords: Adaptative Multi-Agent System · Ensemble learning · Non-linear classification.

1 Introduction

Supervised classification problems have been extensively addressed using machine learning algorithms called classifiers. These problems are generally not linearly separable (i.e. the point clouds of the different classes are not separable by hyperplanes). Moreover, in so-called dynamic problems, new classes appear and disappear and the behavior of individuals evolves in time and space (e.g. management of smart cities and the appearance of new transport modes and behavior). Dynamic problems are often associated with ambient environments in which new devices may appear or disappear dynamically [12, 16]. Many classifiers are themselves composed of multiple linear classifiers (i.e. whose decision function is a linear combination of input variables) that they aggregate.

After positioning *smapy* with respect to machine learning techniques for classifier aggregation in section 2, we detail its operation in section 3. In section 4.2, we show that our MAS is able, thanks to a space exploration by four types of

linear classifiers and cooperation rules, to linear toy classification problems with different levels of linear separability from the literature. **We show through this experiment that it is possible to transform a classification problem into a cooperation problem** through the exploration of the space of input variables. The results presented in section 5 are discussed in section 6.

The main contributions of this paper are:

- *Smapy*, an ensemble Multi-Agent System (MAS) for online non-linear classification based on the AMAS framework
- An experiment on three two-dimensional problems from literature to evaluate the performance of our MAS for different levels of linear separability

2 Related work

In this section, we situate our approach among other machine learning methods based on aggregation of classifiers. After presenting their advantages and weaknesses, we show that the AMAS approach is a constructivist and ensemble learning method. We then detail how our approach meets the objectives of **online** solving of **dynamic problems** through cooperative tiling of the input variable space.

2.1 Aggregation of classifiers

Many non-linear classifiers aggregate several linear classifiers. We distinguish two types of aggregation: connectionist and ensemble approaches. The resulting algorithms can be self-constructing.

Connectionism The connectionist approach can be illustrated by a serial electrical circuit. The input data passes successively through a chain of classifiers and the decision function of the global classifier is a composition of the decision functions of the internal classifiers. In supervised learning, the information on the output goes back up the chain of classifiers in the opposite direction.

The connectionist approach is associated with artificial neural networks (ANN) in which the information of the output data is fed back through the gradient back-propagation algorithm [17]. The internal classifiers composing the neural network are linear. A network is usually composed of several layers (multi-layer perceptron). The multiplication of these layers is at the origin of deep learning.

The main limitation of the connectionist approach is the need for a large number of observations to optimize the weights associated with each internal classifier. In dynamic problems where the behavior of individuals evolves over time, new data are generally not available in sufficient number to update the learning model.

Ensemble The ensemble approach is similar to a parallel electrical circuit. The input data passes through several classifiers in parallel, and the output of the decision function is an aggregation of the outputs of the decision functions of each classifier.

Bagging is a learning technique combining the use of bootstraps and the aggregation of prediction models. The assumption of bagging, inspired by the law of large numbers, is that averaging the predictions of several independent models reduces the variance and thus the error of the global prediction. The Random Forest algorithm [2] uses binary decision trees by adding a random draw of the input variables to be considered for each intermediate model.

Unlike bagging, boosting algorithms build a model sequentially from so-called weak models. At each step, the bad points predicted by the previous model are given a higher weight when training the current model. Adaptive Boosting (AdaBoost) [18] uses binary decision trees with a single node and a single input variable. In Gradient Boosting [11] and eXtreme Gradient Boosting (XGBoost) [5], the weights of the points are no longer incremented but a cost function minimized by gradient descent allows to aggregate intermediate models to the global model.

Constructivism In self-constructing aggregation approach, the graph of internal classifiers (connectionist or ensemble) does not have a fixed structure. This structure is built during the learning process and can evolve over time. When changes to the classifier graph are made by the classifiers themselves, the system is self-organizing.

Algorithms such as self-constructing neural networks [4] lie at the intersection of connectionism and constructivism, because the structure of the layers of neurons is built incrementally.

Context-aware learning is a constructivist and almost ensemble approach. In ELLSA [8], the Context agents can be seen as linear classifiers (although they can only predict one class) with cooperative rules.

The constructivist approach allows us to design self-organizing non-linear classifiers suitable for solving dynamic problems. In this paper, we propose an ensemble context-based learning MAS. In the following paragraphs, we situate our contribution in the field of adaptative MAS.

2.2 Multi-Agent Systems

Multi-agent systems (MAS) are systems involving autonomous entities called agents, capable of communicating in a common environment in which each has its own perceptions and knowledge [10]. Each agent acts locally to get closer to its objectives using its own skills.

Problem solving by MAS is the search for a balance in the interactions between agents. It is not the sum of the individual abilities of the agents, but rather the product of their interactions, which leads to the emergence of new resolution

abilities. This phenomenon of emergence is regularly invoked to justify the use of MAS for the collective resolution of dynamic and complex problems.

Self-organization Self-organization is the capacity of a system to dynamically modify its internal structure without external intervention [14]. Self-organization is said to be strong when the modifications are not the consequence of a centralized decision. In the framework of MAS, the agents can modify the global function of the system by acting on their local environment. This process of adaptation to the dynamics of the problem is similar to machine learning when the agents share a common goal.

The theory of adaptive multi-agent systems (AMAS) [3] proposes a cooperative approach to interactions between agents. The design criteria presented for these interactions guarantee a satisfactory, but not necessarily optimal, result in the resolution of the problem at hand (functional adequacy theorem).

Context learning Context learning consists in exploring the space defined by the input variables of the model using cooperating agents. The AMAS for context learning (AMAS4CL) approach is based on the AMAS theory and more particularly on the Self-Adaptive Context Learning (SACL) [1] paradigm to define the rules of cooperation between agents and proposes a structure composed of several types of agents to explore the space of the problem variables. Algorithms based on the SACL approach are used to solve various problems such as learning by demonstration [20] or Inverse Kinematics [9] in robotics and optimization of the operation of a heat pump [1]. SACL architectures are typically composed of three types of agents:

- Context agents that define hypercubes of the input variable space. When a new point belongs to one of these zones, the corresponding Context agent is said to be activated and proposes a system decision according to its own knowledge
- The Percept agents which retrieve the values of the input variables (sensors) at each iteration and transmit them to the Context agents
- The Head agent which receives the proposals of the activated Context agents and sends them feedbacks from oracle

A context learning MAS has two modes of operation: exploration (learning) which consists in tiling the input variable space by instantiating and arranging the Context agents thanks to the feedback from the Head agent, and exploitation (prediction) which consists in taking a decision without updating the system. The functioning of this architecture is presented in Fig. 1 for the cooperative case in the exploration phase (i.e., the optimal functioning case). When the behavior of the system is not optimal with respect to the user’s objectives, the situation is said to be non-cooperative (NCS). The system must adapt to maximize the cooperation between agents to return to the cooperative state. In Context Learning, this cooperation is expressed in the sizes (i.e., the dimensions of the hypercubes), positions and knowledge of the Context agents.

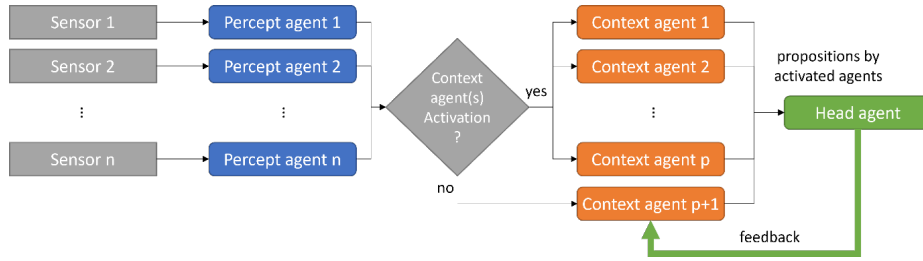


Fig. 1. Cooperative operation of a SACL architecture in exploration phase

The SACL approach is adapted to dynamic problems (i.e. the system adapts to changes in the distribution of the input data over time) and supports online learning. Moreover, the position and size of the agents in the space of the input variables provide additional information about the phenomenon under study and a geometric interpretation.

3 Smapy

In our contribution, we provide each Context agent with an internal machine learning model, linear or not, with the only requirement to support online learning. Each internal model is trained on the points that have activated the corresponding Context agent and thus constitutes a local modeling (in the sense of the space of input variables) of the underlying function of the problem to solve. This context learning MAS, *smapy*, has been implemented in python for the industrial needs of the research project.

3.1 General principle

Like other SACL type architectures, *smapy* has two modes of operation:

- The exploration during which the mapping of the space of the input variables is modified according to new available labelled observations
- The operation during which the system uses its coverage of the space of input variables to classify a new point

In both cases, the operation of the system is iterative, and each cycle starts with a new observation. During the exploration, the activated Context agents update their internal model with the last observation after they have proposed an output class to the Head agent and received feedback (positive or not). The feedback received by a Context agent allows it to update its perception of itself within the collective through a performance metric explained in the section 3.3. It also allows him to know if he has a non-cooperative behavior with respect to the objective of the system and, if necessary, to act on itself or its neighbors to return to a cooperative state (c.f. section 3.4).

3.2 Agents

In this section, we present the three types of agents involved in our SACL architecture, whose relationships have been described in Fig. 1.

Percept The p Percept agents collect the values of the p input variables of each new observation and pass them to the Context agents. They also store the observed extrema for each variable.

Context A Context agent l defines a hypercube in the p -dimensional space of input variables. For each dimension j , it has two parameters $r_{l,j,0}^t$ and $r_{l,j,1}^t$ that define the lower and upper bounds of an activation interval at iteration t . The agent can compute at any time $v_l(T)$, the volume of its activation hypercube at iteration T , according to the following formula:

$$v_l(T) = \prod_{j=1}^p (r_{l,j,1}^t - r_{l,j,0}^t) \quad (1)$$

The Context agent also has a confidence level $c_l(T)$ at iteration T , depending on its history \mathcal{H}_l^T (set of its activation cycles since its creation), its class proposals \hat{y}_l^t for observations y^t on this history, and two external parameters F_+ and F_- that respectively weight the positive and negative feedbacks of the agent Head:

$$c_l(T) = \sum_{t \in \mathcal{H}_l^T} (F_+ * \mathbb{K}_{\hat{y}_l^t = y^t} - F_- * \mathbb{K}_{\hat{y}_l^t \neq y^t}) \quad (2)$$

From its two terms, we define the score $s_l(T)$ of a Context agent at iteration T using a normalization function N_c which is an external parameter of *smapy*:

$$s_l(T) = N_c \circ c_l(T) \quad (3)$$

Finally, the Context agent has an internal classification model learned from the observations that activated it. The python implementation of *smapy* makes it possible to use models in the *scikit-learn* fashion if they support online learning to adapt the agent to new observations. For the rest of this paper, we define several properties of Context agents:

Definition 1. (Expansion/retraction) A Context agent expands (resp. retracts) by a factor α when it increases (resp. decreases) its boundaries to multiply its volume by $1 + \alpha$ (resp. $1 - \alpha$).

Definition 2. (Push) A Context agent l_1 pushes a Context agent l_2 when l_2 retracts so that the previous intersection of l_1 and l_2 is completely outside l_2 (and thus contained only within l_1).

Definition 3. (Absorption) A Context agent l_1 absorbs a Context agent l_2 when l_1 expands to completely contain the area covered by l_2 and the agent l_2 is destroyed.

Definition 4. (Point exclusion) *A Context agent l_1 excludes an observation y when l_1 retracts so that y ends up outside l_1 . Point exclusion is controlled by an external Boolean parameter E .*

Definition 5. (Overlapping index) *The overlapping index o_{l_1, l_2} is the ratio of the volume of the intersection of two Context agents l_1 and l_2 to the minimum of the volumes of these agents:*

$$o_{l_1, l_2} = o_{l_2, l_1} = \frac{v_{l_1 \cap l_2}}{\min(v_{l_1}, v_{l_2})} \quad (4)$$

Head The Head agent supervises the cooperation of the Context agents. At each iteration, it selects the class proposed by the activated Context agent with the highest score (and proceeds by vote in case of a tie) and sends feedbacks to all the agents activated during the exploration phase (c.f. section 3.3). The Head agent can also create new Context agents in case of system incompetence (c.f. section 3.4).

3.3 Feedback

When the Context agents are activated, they propose a prediction to the Head agent. The latter selects the prediction of the agent with the highest score. During the exploration phase (learning), the Head agent sends feedbacks to the Context agents which have proposed a prediction:

- If the prediction is good (with respect to the label of the new point), then the confidence of the context agent increases by F_+ and it expands by a factor α (external parameter)
- If the prediction is bad, then the confidence of the context agent decreases by F_- . If point exclusion is allowed (i.e., E is true), then the context agent excludes the new point. Otherwise, the Context agent's local model is fine-tuned with the new point (in the sense of online learning), and it retracts by a factor α

3.4 Non-cooperative situations

The objective of AMAS is to transform the initial problem into a problem of cooperation between agents. Non-cooperative situations (NCS) are states during which the behavior of the system must evolve to reach the goal set by the user. In context learning, this results in the rearrangement of the Context agents to improve the tiling of the input variable space. In this section, we present and schematize in Fig. 2 the three types of NCS that can occur during the operation of *smapy* and their resolution.

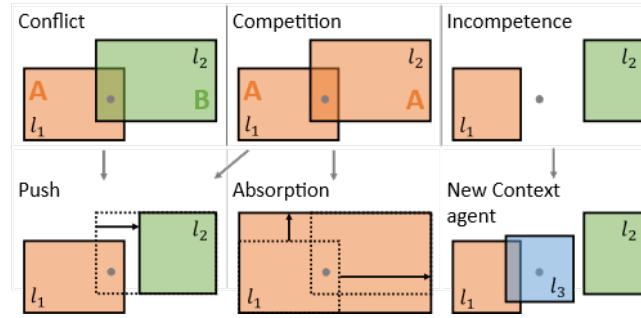


Fig. 2. Schematic of NCS (top row) and their resolution (bottom row) for Context agents l_1 , l_2 and l_3 predicting classes A and B

Incompetence Incompetence occurs when no Context agent has been activated:

- Exploration: a new Context agent is created around the new point and any NCS generated are resolved. The initial radius of the new agent is controlled by an external parameter R
- Exploitation: the closest Context agent to the new point (in the sense of the Euclidean distance between the point and the agent’s boundary) proposes its prediction

Competition Competition occurs during the exploration phase when two activated Context agents propose the same prediction (in this case, the same class):

- If an overlapping threshold is defined through the external parameter O , and if the overlapping index of the two agents is greater than this threshold, the agent with the higher score absorbs the other
- Otherwise, the Context agent with the higher score pushes the other agent

Conflict Conflict occurs during the exploration phase when two activated Context agents propose different predictions. The agent with the higher score then pushes the other agent.

4 Comparison of linear classifiers alone with context learning approach

In this section, we present the experiment of comparing four types of linear classifiers and instances of *smapy* with these same models inside Context agents. The linear classifiers used in this paper are :

- Logistic regression (ridge [13], LASSO [19] or ElasticNet [21])
- Linear Support Vector Machine (SVM) [7]

- Passive Aggressive algorithms [6] (PA-I and PA-II)

The motivation of this experiment is to verify if the transformation of a classification problem into a multi-agent cooperation problem allows to improve the performances of the latter.

4.1 Input data

The experiment is conducted on three two-dimensional binary classification toy datasets included in the scikit-learn library [15] and regularly used for model comparison purposes:

- Moons: Two interleaved point clouds explained by the two variables (`noise=0.3`)
- Circles : A circular point cloud embedded in another ring-shaped cloud (`noise=0.2, factor=0.5`)
- Linearly separable : Two point clouds with a linear boundary explained by only one of the two variables

Each data set contains 100 points. They are centered and reduced using the scikit-learn `StandardScaler` model before learning.

4.2 Experimental protocol

Table 1. List of value grids for the search of optimal combinations of parameters of the studied linear models (scikit-learn implementation)

Parameter	Grid of values		
	LOGIT & LINEAR SVM		
<code>alpha</code>	0.0001	0.001	0.01
<code>penalty</code>	l_1	l_2	ElasticNet
	PA-I & PA-II		
<code>c</code>	0.5	1.0	2.0

Table 2. List of value grids for finding the optimal combinations *smapy* parameters

Parameter	Grid of values		
R	0.1	0.2	0.5
O	0.2	0.5	
E	False	True	
N_c	Sigmoid		
α	0.0	0.1	0.2
F_+	1.0		
F_-	0.5	1.0	2.0

Step 1 First, we search for the optimal combination of parameters for each of the four linear classifiers among a grid of parameters presented in the table 1 using a five-fold cross validation.

Step 2 Once these combinations are obtained, we search for the optimal combination of *smapy* parameters for each linear classifier among a grid of parameters presented in the table 2 with a five-fold cross validation. The Context agents of the *smapy* instances have as internal model the corresponding linear model, trained with the parameters of its previously obtained optimal combination.

This protocol is repeated for each dataset to obtain 12 *smapy* instances and 12 corresponding linear classifier instances, all of which have been optimized by cross-validation. The optimized *smapy* instances are then compared with the linear classifiers using two evaluation metrics:

- Classification accuracy (multi-class) averaged over the five iterations of cross-validation (step 1 for linear models, step 2 for *smapy*)
- Decision boundaries of the models (linear or *smapy*) trained with the best parameter combinations obtained by cross-validation

5 Results

In this section, we present comparative results between linear classifiers alone and *smapy* instances according to the two metrics introduced previously (section 4.2).

5.1 Classification accuracy

Table 3. Comparison of the classification accuracies obtained for each classifier alone or inside *smapy*

	Moons		Circles		Linearly Separable	
	Alone	MAS	Alone	MAS	Alone	MAS
Logit	0.83	0.89	0.49	0.83	0.92	0.91
Linear SVM	0.86	0.89	0.53	0.83	0.90	0.91
PA-I	0.82	0.89	0.53	0.83	0.89	0.90
PA-II	0.84	0.87	0.53	0.83	0.89	0.89

Table 3 shows no significant difference in accuracy after switching to MAS with the dataset *Linearly separable*. The linear classifiers already achieve a high score despite the noise in the data.

For the other two datasets, we observe an improvement of the accuracy with the *smapy* instances. In particular, we observe a $\sim 30\%$ improvement for the dataset *Circles*, for which the linear classifiers alone give a result close to chance (50%). The poor performance of these classifiers is explained by the low linear separability of the dataset. However, we see that the integration of these classifiers in an SMA allows to approach the quadratic boundary of the dataset.

The accuracies obtained on the dataset *Moons* (intermediate case in terms of linear separability) are slightly better with the MAS approach, but the linear classifiers alone allow to obtain satisfactory scores.

5.2 Decision boundaries

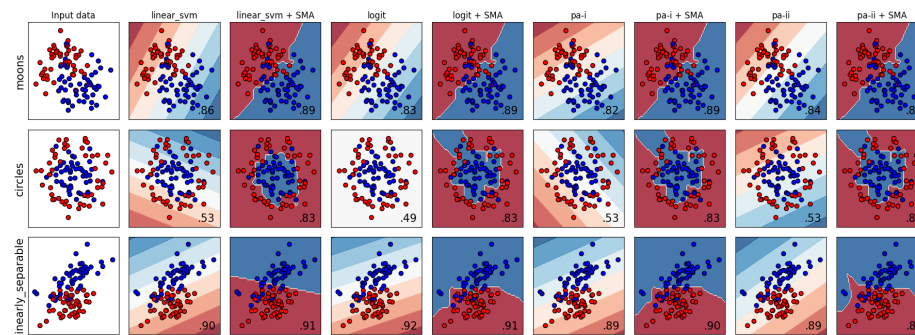


Fig. 3. Decision boundaries and classification accuracy obtained for each dataset (rows) and for each linear classifier alone or in a *smapy* instance (columns)

Fig. 3 shows that linear classifiers alone give linear boundaries that are suitable for the classification problem for the datasets *Moons* and *Linearly separable*. The MAS approach on these two datasets reproduces this linear behavior in the boundaries.

On the other hand, with the dataset *Circles*, the linear models alone are unable to separate the two enclosed point clouds, contrary to the MAS approach.

6 Discussion

Using a MAS approach, the initial classification problem is solved locally at the Context agent level. Thus, even if the agents have linear internal classifiers, they have positioned and sized each other in such a way as to locally approximate a non-linear boundary thanks to the various cooperation mechanisms presented in the section 3.

However, the Context agents may have over-specialized locally by observing homogeneous groups of individuals (in the sense of the class). The existence of the point exclusion mechanism, although often selected by cross validation, tends to reinforce this over-specialization of agents by excluding new class points from their activation zones.

Nevertheless, the ideal behavior sought for *smapy* is to build Context agents that cover homogeneous areas of the explored input variable space, notably for reasons of explicability. There is therefore a trade-off between the geometric

interpretability of the layout of the Context agents and the generalization of the system to dynamic problems in which new classes may appear.

7 Conclusion

Our contribution lies at the intersection of the constructivist SACL pattern and ensemble learning methods. Our contribution provides each Context agent with an internal supervised classification model, as well as rules for cooperation with other agents. By choosing linear models for the Context agents, we show that it is possible to simplify a non-linear classification problem by transforming it into a local cooperation problem within a context learning MAS. Our experimental methodology allows us to observe a significant improvement of the classification accuracy on non linearly separable datasets.

The next step is the use of *smapy* for dynamic real-world problems such as smart city management with the guarantee of an interpretable prediction compared to other state-of-the-art algorithms.

Our main research direction on *smapy* is the possibility to use different algorithms in the internal models of the Context agents. The idea is to exploit the strengths and weaknesses of different known algorithms to optimize prediction quality at specific locations in the space where certain models perform best.

Finally, improvements in the operating rules of *smapy* are needed to avoid over-specialization of the Context agents while maintaining the explicability and stability of the system. For this purpose, the "severity" of the NCS correction mechanisms can evolve according to the convergence of the agents' layout towards a supposedly optimal layout.

Acknowledgements We thank the National Association for Research and Technology (ANRT) for the CIFRE funding of the thesis project in partnership with the Institut de Recherche en Informatique de Toulouse (IRIT) and the company Citec Ingénieurs Conseil. We also thank all the reviewers for their help and advice.

References

1. Boes, J., Nigon, J., Verstaevel, N., Gleizes, M.P., Migeon, F.: The self-adaptive context learning pattern: Overview and proposal. In: International and Interdisciplinary Conference on Modeling and Using Context. pp. 91–104. Springer (2015)
2. Breiman, L.: Random forests. *Machine learning* **45**(1), 5–32 (2001)
3. Capera, D., Georgé, J.P., Gleizes, M.P., Glize, P.: The amas theory for complex problem solving based on self-organizing cooperative agents. In: WET ICE 2003. Proceedings. Twelfth IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprises, 2003. pp. 383–388. IEEE (2003)
4. Carmo Nicoletti, M.d., Bertini, J.R., Elizondo, D., Franco, L., Jerez, J.M.: Constructive neural network algorithms for feedforward architectures suitable for classification tasks. In: Constructive neural networks, pp. 1–23. Springer (2009)

5. Chen, T., Guestrin, C.: Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. pp. 785–794 (2016)
6. Crammer, K., Dekel, O., Keshet, J., Shalev-Shwartz, S., Singer, Y.: Online passive aggressive algorithms. *Journal of Machine Learning Research* **7**, 551–585 (2006)
7. Cristianini, N., Shawe-Taylor, J., et al.: An introduction to support vector machines and other kernel-based learning methods. Cambridge university press (2000)
8. Dato, B.: Lifelong Learning by Endogenous Feedback, application to a Robotic System. Ph.D. thesis, Université Toulouse 3-Paul Sabatier (2021)
9. Dato, B., Gleizes, M.P., Migeon, F.: Cooperative neighborhood learning: Application to robotic inverse model. In: 13th International Conference on Agents and Artificial Intelligence (ICAART 2021) (2021)
10. Ferber, J., Weiss, G.: Multi-agent systems: an introduction to distributed artificial intelligence, vol. 1. Addison-Wesley Reading (1999)
11. Friedman, J.H.: Stochastic gradient boosting. *Computational statistics & data analysis* **38**(4), 367–378 (2002)
12. Guivarch, V., Camps, V., Péninou, A., Stuker, S.: Dynamic filtering of useless data in an adaptive multi-agent system: evaluation in the ambient domain. In: International Conference on Practical Applications of Agents and Multi-Agent Systems. pp. 110–121. Springer (2013)
13. Hoerl, A.E., Kennard, R.W.: Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* **12**(1), 55–67 (1970)
14. Marzo Serugendo, G.D., Gleizes, M.P., Karageorgos, A.: History and definitions. In: Self-organising software, pp. 33–74. Springer (2011)
15. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E.: Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research* **12**, 2825–2830 (2011)
16. Perles, A., Crasnier, F., Georgé, J.P.: Amak-a framework for developing robust and open adaptive multi-agent systems. In: International Conference on Practical Applications of Agents and Multi-Agent Systems. pp. 468–479. Springer (2018)
17. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning representations by back-propagating errors. *nature* **323**(6088), 533–536 (1986)
18. Schapire, R., Freund, Y., et al.: A decision-theoretic generalization of on-line learning and an application to boosting. In: Second European Conference on Computational Learning Theory. pp. 23–37 (1995)
19. Tibshirani, R.: Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)* **58**(1), 267–288 (1996)
20. Verstaavel, N., Régis, C., Gleizes, M.P., Robert, F.: Principles and experimentations of self-organizing embedded agents allowing learning from demonstration in ambient robotic. *Procedia Computer Science* **52**, 194–201 (2015)
21. Zou, H., Hastie, T.: Regularization and variable selection via the elastic net. *Journal of the royal statistical society: series B (statistical methodology)* **67**(2), 301–320 (2005)