MALEF

A Framework for Agent-Based Distributed Machine Learning and Data Mining

Jan Tožička[1]
Michael Rovatsos[2]
Michal Pěchouček[1]

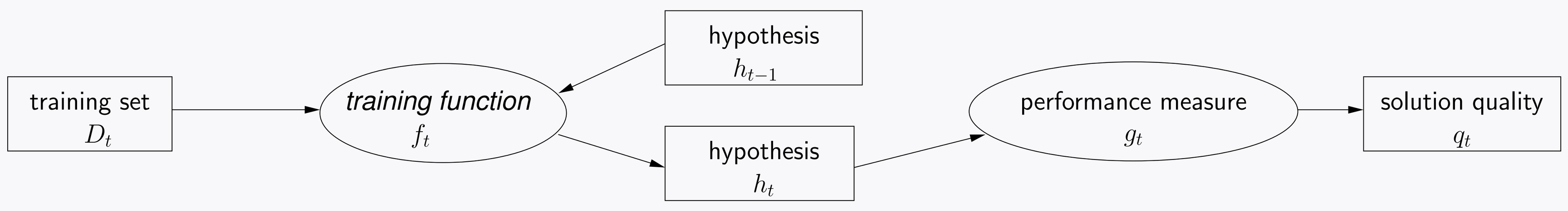
[1] Agent Technology Group, Gerstner Laboratory,
Czech Technical University in Prague
tozicka,pechouc@labe.felk.cvut.cz

[2] School of Informatics, The University of Edinburgh,
Edinburgh, United Kingdom
mrovatso@inf.ed.ac.uk

Objective: ***To devise a sufficiently general, abstract view of describing autonomous learning processes in order to be able to utilise the whole range of methods for (i) rational reasoning and (ii) communication and coordination offered by agent technology so as to build effective distributed learning systems.***

Generic Model of Learning Process

A single iteration of a learning process can be seen as an agent reasoning cycle that has the following structure:



Can we exploit this to investigate rich forms of interaction between learners in a distributed machine learning system?

Learner Coordination

Integration Matrix

The generic learner model allows us to look at the different possibilities for information exchange among learners:

i ^{j}	D_j	H_j	f_j	g_j	h_j
D_i	$p_1^{D\to D}(D_i, D_j)$ ⋮ $p_{k_{D\to D}}^{D\to D}(D_i, D_j)$	…	…	n/a	…
H_i	⋮	⋱		n/a	
f_i	⋮		⋱	n/a	
g_i	⋮			n/a	$p_1^{g\to h}(g_i, h_j)$ ⋮ $p_{k_{g\to h}}^{g\to h}(g_i, h_j)$
h_i	⋮			n/a	⋱

Knowledge exchange problem: **what** knowledge to exchange, **when** to exchange the knowledge, and **how** to use received knowledge

Evaluation of Example Merging Operators

So far, we have experimented with merging the models/hypotheses of different learners using a contract-net-style approach.
Let h_j the receiver's own model h_i the provider's model in a clustering scenario:

:: $p^{h\to h}(h_i, h_j)$:

- **m-join:** The m best clusters (in terms of coverage of D_j) from hypothesis h_i are appended to h_j.
- **m-select:** The set of the m best clusters (in terms of coverage of D_j) from the union $h_i \cup h_j$ is chosen as a new model. (Unlike m-join this method does not prefer own clusters over others'.)

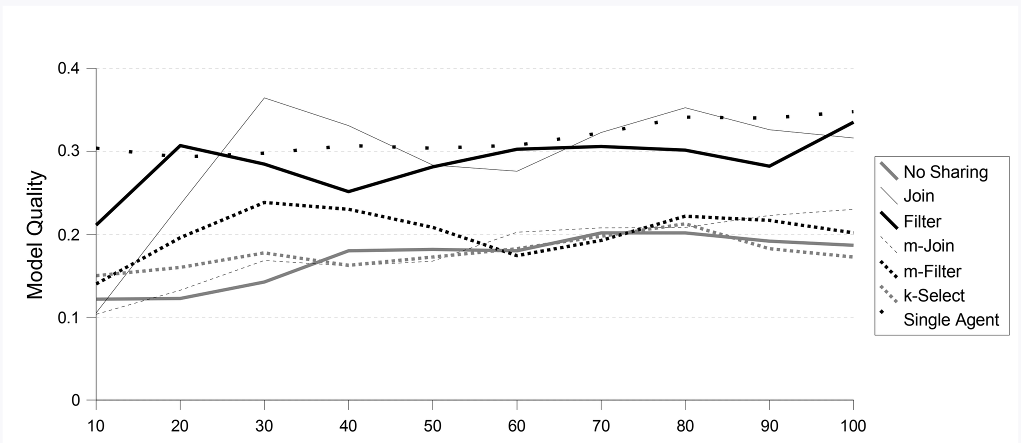
:: $p^{h\to D}(h_i, D_j)$:

- **m-filter:** The m best clusters (as above) from h_i are identified and appended to a new model formed by using those samples not covered by these clusters applying the own learning algorithm f_j.

Experimental Results

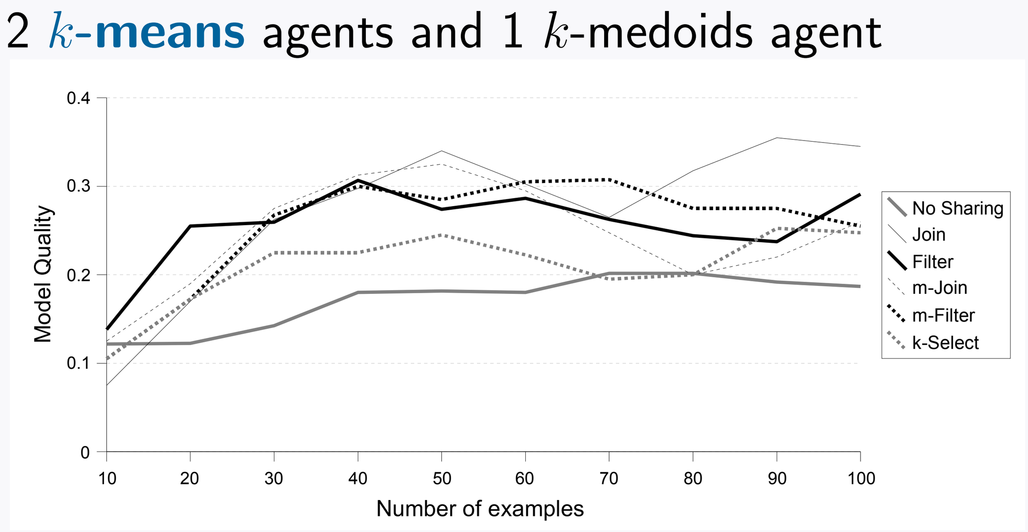
Homogeneous Learners

3 **k-means** agents



Heterogeneous Learners

2 **k-means** agents and 1 k-medoids agent



Performance: Speed Up

The **m-filter** operation, decreases the number of learning samples and thus can **speed up** the learning process.

	k-means	k-medoids
filtering	30-40 %	10-20 %
m-filtering	20-30 %	5-15 %

Conclusion

Even a very simplistic application of **MALEF** architecture has proven capable of **improving the performance** of individual learning agents.
In the future we want to explore other types of combining the different elements of different learners through elaborate communication and reasoning methods.

Acknowledgement

We gratefully acknowledge the support of the presented research by Army Research Laboratory project N62558-03-0819 and Office for Naval Research project N00014-06-1-0232.