THE MASH PROJECT



Validation in Statistics and ML WIAS, Berlin

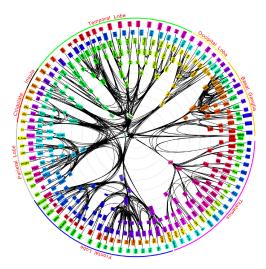






INTRODUCTION COMPLEXITY IN MACHINE LEARNING

Artificial learning systems remain extremely simple compared to their biological counterparts.



(Macaque brain long-distance network, Modha and Singh, 2009)

The MASH project

In practice, increasing complexity of ML by simply combining different modalities or prediction methods is very efficient:

- Different modalities tend to catch *complementary information*.
- Different prediction methods tend to be wrong differently.

In both case, one ends up combining "independent" variables, and the overall prediction gets better.

François Fleuret

The MASH project

3 / 29

NTRODUCTION COMPLEXITY IN MACHINE LEARNING

The Netflix Prize started in October 2, 2006, and was won in September 2008 by "BellKor's Pragmatic Chaos".

... we use a set of diverse state-of-the-art collaborative filtering (CF) algorithms, which include: SVD, Neighborhood Based Approaches, Restricted Boltzmann Machine, Asymmetric Factor Model and Global Effects. We show that linearly combining (blending) a set of CF algorithms increases the accuracy and outperforms any single CF algorithm.

(Jahrer et al., 2010)

NTRODUCTION COMPLEXITY IN MACHINE LEARNING

Flower classification

Sin	gle feature	s	Combinat	ion metho	ds
Method	Accuracy	Time	Method	Accuracy	Time
Colour	$\textbf{60.9} \pm \textbf{2.1}$	3	Product	85.5 ± 1.2	2
Shape	$\textbf{70.2} \pm \textbf{1.3}$	4	Averaging	84.9 ± 1.9	10
Texture	63.7 ± 2.7	3	CG-Boost	84.8 ± 2.2	1225
HOG	58.5 ± 4.5	4	MKL (SILP)	85.2 ± 1.5	97
HSV	61.3 ± 0.7	3	MKL (Simple)	85.2 ± 1.5	152
siftint	$\textbf{70.6} \pm \textbf{1.6}$	4	LP- β	85.5 ± 3.0	80
siftbdy	59.4 ± 3.3	5	LP-B	85.4 ± 2.4	98

(Gehler and Nowozin, 2009)

The *learning method* does not matter much.

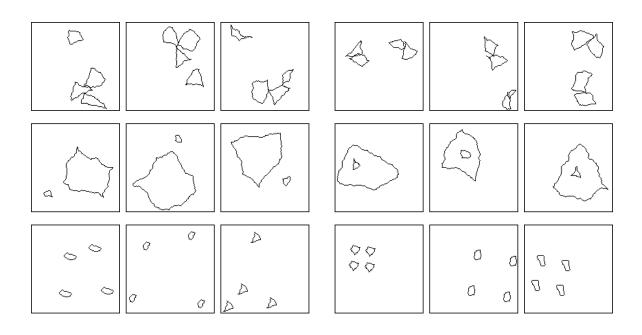
François Fleuret

The MASH project

5 / 29

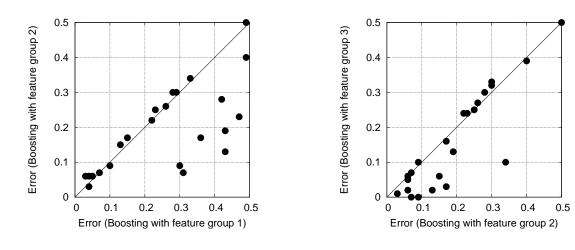
COMPLEXITY IN MACHINE LEARNING

The Synthetic Visual Reasoning Test.



NTRODUCTION COMPLEXITY IN MACHINE LEARNING

Results on the Synthetic Visual Reasoning Test.



Group 1: Pixel counting Group 2: Group 1 + Edge-like Group 3: Group 2 + Fourier-like

François Fleuret

The MASH project

7 / 29

NTRODUCTION THE FOG OF MACHINE LEARNING

So, complex learning systems are worth investigating. However, ML development faces specific difficulties as an engineering task:

- Specifications involve a very complex object (data set, real-world POMDP)
- Limited understanding beyond rough behaviors: Convergence, over-fitting, cost, some invariance.
- Resulting algorithms combine very large numbers of (simple) cues. The emerging behavior is of a different nature.

We often have no idea why it truly works (buggy code sometime works as well ...)

In practice, using ML for applications is a meta-learning algorithm:

Developers go back and forth between identifying mistakes (*High-frequencies patterns generate false alarms!*), fixing them (*Let's add features to detect high frequency blobs!*), repeat (*we do not detect bald people's faces! Let's add features to pick roundish shapes!*), and repeat (*now we are over-fitting! Let's add a L¹ penalty!*)

This is similar to Boosting or SVM: At any moment, the most severe errors drive greedy changes in the constructed predictor.

Human are super-optimizers seeing (a bit) more than the gradient.

François Fleuret

The MASH project

9 / 29

THE MASH PROJECT MOTIVATION

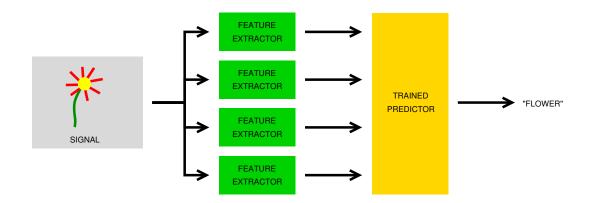
The MASH project is a three-year European research initiative motivated by these observations:

- Machine learning lacks tools to rationalize the design of very complex architectures.
- Combining multiple feature extractors and prediction methods improves performances.
- Internet based collaborative tools allow large teams of individuals to work together.

We want to create new tools for designing complex learning systems in a collaborative manner.

THE MASH PROJECT FEATURE EXTRACTION

The project focuses on standard combinations of feature extractors and ML methods.



Researchers in the project develop new ML and analysis tools, while *external contributors* design feature extractors.

François Fleuret

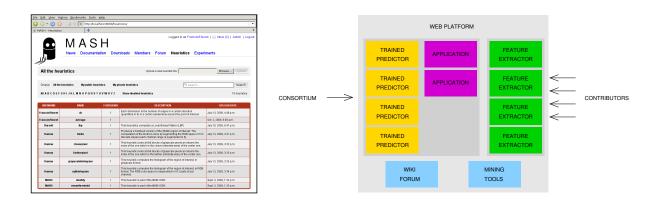
The MASH project

11 / 29

THE MASH PROJECT OVERALL ARCHITECTURE

Research is organized through a web collaborative platform at

http://mash-project.eu





Vision tasks



Image classification



Object detection

François Fleuret

The MASH project

13 / 29

THE MASH PROJECT APPLICATION

Goal-planning



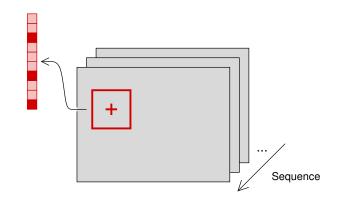
Simulated environment



Robotic arm



We define the concept of *Heuristic*, a feature extractor with a persistent state.



Contributions are C++ sources implementing such a heuristic.

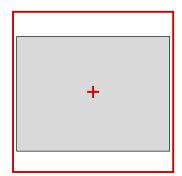
François Fleuret

The MASH project

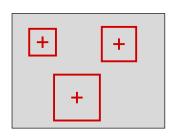
15 / 29

THE MASH PROJECT

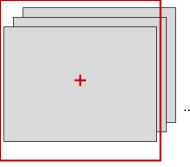
HEURISTICS FOR APPLICATIONS



Classification



Detection



Goal-planning

THE MASH PROJECT NOTION OF HEURISTIC

If we ignore the persistent state, a heuristic is a mapping

```
H: [0,1]^{3WH} \times \{1,\ldots,W\} \times \{1,\ldots,H\} \to \mathbb{R}^{D}
```

```
void init();
unsigned int dim();
void prepareForImage();
void finishForImage();
void prepareForCoordinates();
void finishForCoordinates();
scalar_t computeFeature(unsigned int feature_index);
```

François Fleuret

The MASH project

17 / 29

COLLABORATIVE PLATFORM CONTRIBUTOR TOOLS

The tools to help contributors include three main components:

- Documentation and social tools: Wiki, screencasts, forum, and private messaging.
- A multi-platform Software Development Kit.
- On-line development tools:
 - **Heuristic repository:** Hosts multiple versions of each heuristics, private or public. Keeps track of the phylogeny.
 - **Private experiments:** Combine a limited number of heuristics, and reduced settings for the database and predictors.
 - **Public experiments:** Defined by the consortium, combine many heuristics and full-scale data-sets.

COLLABORATIVE PLATFORM CONTRIBUTOR TOOLS

The interaction between a contributor and the MASH system should follow roughly the following steps:

- 1. Look at the existing heuristics, and the current performance of the system on the public experiments.
- 2. Download and improve one heuristic, or write one from scratch
- 3. Upload it, the platform checks it works, runs a series of tests (two settings, ten runs on each at the moment), and provide a ranking.
- 4. Run additional private experiments to assess more specific strengths and weaknesses of the new contribution, and its complementarity with others.

François Fleuret

The MASH project

19/29

COLLABORATIVE PLATFORM

SCREENSHOTS





COLLABORATIVE PLATFORM SCREENSHOTS

MASH News Documentation Downloads Members		uret 🔤 Inbox (0) Admin Logout Experiments			MAS News Document		vnloads l	Membe	Logged in as FrancoisFlet	uret ⊠ Inbox (0) Admin Log Experiments Servers
	Q. Search	Search Advanced search		Members					Search	Search
23User Control Panel (0 new messages) - View your posts Moderator Control Panel]		QFAQ		AII A B C D E F	GHIJKLMNOPQRST	U V W X Y Z #				36 users
rancoisfleuret/zk				USERNAME	E-MAIL	GROUP	HEURISTICS	POSTS	WEBSITE, LOCATION	JOINED
POSTREPLY # 9. Search this topic. Search		2 posts • Page 1 of 1		Kanma	philip.abbet@idiap.ch	Core team	5	3	http://www.idiap.ch/~pabbet Idiap Research Institute, Switzerland	June 26, 2009, 1:50 p.m.
		-	_	ftarsett	flavio.tarsefi@idiap.ch	Contributor	2	0	http://www.idiap.ch Idiap.Research Institute, Switzerland	June 26, 2009, 2:05 p.m.
Dby MASH > March 24th, 2010, 9:36 am			ш.	FrancoisReuret	francois.fleuret@idiap.ch	Core team	6	1	http://www.idiap.ch/-fleuret/ Idiap Research Institute, Switzerland	June 26, 2009, 2:17 p.m.
'his is the official topic of the francoisfleuret/zk heuristic.		Ru.		jojo	marietho@idiap.ch	Contributor	0	0		June 26, 2009, 2:48 p.m.
		MASH		monay	florent.monay@gmail.com	Contributor	0	0		June 26, 2009, 2:57 p.m.
		The System		omasson	olivier.masson@idiap.ch	Contributor	0	0		July 6, 2009, 2:21 p.m.
		Posts: 12 Joined: September 3rd, 2009, 12:28 pm		PuckCh	olivier.bomet@idiap.ch	Contributor	0	0	http://puck.ch/ Idiap Research Institute, Switzerland	July 6, 2009, 3:42 p.m.
		8** 3		ananchen	alexandre.nanchen@idiap.ch	Core team	0	0		July 13, 2009, 2:27 p.m.
		<u>a</u> .		MASH	mash@mash-ml.eu	The System	2	12	http://www.mash-project.eu	Sept. 3, 2009, 1:27 p.m.
				cdubout	charles.dubout@idiap.ch	Core team	5	0	Idiap Research Institute, Switzerland	Oct. 12, 2009, 7:27 p.m.
le: francoisfleuret/zk	*EDIT 🕷 V V (fouote)			smarcel	marcel@idiap.ch	Contributor	0	0		Dec. 16, 2009, 9:05 a.m.

François Fleuret

The MASH project

21 / 29

COLLABORATIVE PLATFORM SCREENSHOTS

0	MAS News Documenta		Logged in as FrancoisFleure	t 🔤 Inbox (0) Admin Logos	Logged in as Francost-Reuret a thook (3) News Documentation Downloads Members Forum Heuristics Experiments Se	
All the he			Uplead a new heutistic file whether bearistics	Browse Upload Search 28 heuristics	Addrer: Francoh/Terent Upland date: Jane 21, 2010, 159 p.m. Accessibility: Pale: Status: Enable Summary: Computes proportion of edges over rectangles picked at random. Bandatate: Computers (Configuration I Training ency (sof)) Tree ency (sof) Reak	/ 1
USERNAME	NAME	# VERSIONS	DESCRIPTION	UPLOAD DATE	 adaboost-catech 65.53% (2.1103) 71.0% (1.375) 2 perceptron-mist 0.0% (0.023) 4	
cdubout	hough	1	Computes the linear Hough transform of the region of interest to detect lines.	April 20, 2010, 12:35 p.m.	 All configurations 2	
cdubout	fourier	1	Fourier transform heuristic. Computes the 2D FFT of the region of interest to convert it to the frequency domain.	April 20, 2010, 12:39 p.m.		
cdubout	haar	1	Hear transform heuristic. Computes the 2D hear transform of the region of interest (all levels).	April 20, 2010, 8:18 p.m.	 Source code Previous versions (2) Derived heuristics (6)	
cdubout	identity	1	Same as the mash/identity heuristic except that it returns RGB values instead of grayscale.	April 20, 2010, 8:24 p.m.	 VERSION ACCESSIBILITY UPLOAD DATE	TOOLS
cdubout	hog	1	Histogram of oriented gradients (orientation discretized in 12 bins) taken at random positions and scales. Strongly inspired from francoisfieuretzk_v2.	May 3, 2010, 12:28 p.m.	1 (diff) Public March 24, 2010, 9:29 a.m. 2 (diff) Public April 24, 3010, 9:30 a.m.	-
FrancoisFleuret	average	1	Return the average level of the pixel gray levels.	Oct. 2, 2009, 8:08 a.m.		
FrancoisFleuret	boxedaverages	2	Computes the average gray level over rectangular windows picked at random.	April 24, 2010, 9:37 a.m.		

	IVI News	A S		Heuristics Experiments	Servers
My public h	neuristic	s	Upload a new heuristic file	B	rowse Upload
			uristics My private heuristics	9 Search	Search
Display: All the h	euristics	My public her	ansucs my private neuristics	v 568 GT	orditeit
				UPLOAD DATE	5 heuristics
ANABCDEFG	HIJKLM	NOPQRS	T U V W X Y Z Show disabled heuristics		5 heuristics
AII A B C D E F G NAME	HIJKL M	N O P Q R S	T U V W X Y Z Show disabled heuristics DESCRIPTION	UPLOAD DATE Oct. 2, 2009, 8.08 a.m.	5 heuristics TOOLS
All A B C D E F G NAME average	HIJKL M 2 1995	N O P Q R S T # VERSIONS	T U V W X Y Z Show disabled heuristics	UPL OAD DATE Oct. 2, 2009, 8:08 a.m. ked at April 24, 2010, 9:37 a.m.	5 heuristics TOOLS
All A B C D E F G NAME average boxedavera	HIJKL M 2 1995	N O P Q R S # VERSIONS 1 2	TUVWXYZ Show disabled heartstics DESCREPTION Result the average level of the point gray levels. Computes the average gray level over restangular windows plot random. For each pair, computes 8 Chamier distances to sam and vers	UPLOAD DATE Oct 2, 2009, 809 a.m. Aed at April 24, 2010, 9:37 a.m. I ⁹ June 21, 2010, 207 p.m.	5 heuristics

COLLABORATIVE PLATFORM SCREENSHOTS

MASH	embers Forum Heu	ristics Experime	nts Servers		Schedule a private experiment
				Name:	A test with MNIST The name of the experiment (leave empty for an auto-generated one)
ivate experiments of FrancoisFleuret + Schedu	ale a new consortium experir	nent 🕂 Schedule a r	new private experiment	Database:	mnist Contains Images of 60000 handwriten digits, from 0 to 9
play. The public experiments The consortium experiments My pr	ivate experiments	9. Search	Search	Labels:	Selected: 0/10 [select all]
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z			8 experiments		2 3 4
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z NAME	CREATION DATE	START DATE	8 experiments DURATION		2 3 4 5 6
	CREATION DATE Aug. 3, 2010, 9:34 a.m.	START DATE Scheduled			2 3 4 5 6 7 7 7
NAME					2 3 4 5 5 6 7 8 8 9 8 9 9
NAME classification®	Aug. 3, 2010, 9:34 a.m.	Scheduled	DURATION 42 minutes, 55 seconds 2 hours, 42 minutes, 59		a a b b b c c c c c c c c c c c c c c c
MANE classification® classification?	Aug. 3, 2010, 9:34 a.m. July 26, 2010, 9:34 p.m.	Scheduled July 26, 2010, 9:34 p.m.	DURATION 42 minutes, 55 seconds	Training ratio:	0.90
NANE Classification 9 classification 7 classification 5	Aug. 3, 2010, 9:34 a.m. July 26, 2010, 9:34 p.m. July 26, 2010, 5:18 p.m.	Scheduled July 26, 2010, 9:34 p.m. July 26, 2010, 5:19 p.m.	DURATION 42 minutes, 55 seconds 2 hours, 42 minutes, 59 seconds		0.90 Ratio of the objects used for training (between 0.0 and 1.0)
NANE classification® classification7 classification6 classification5	Aug. 3, 2010, 9:34 a.m. July 26, 2010, 9:34 p.m. July 26, 2010, 5:18 p.m. July 14, 2010, 11:19 p.m.	Scheduled July 26, 2010, 9:34 p.m. July 26, 2010, 5:19 p.m. July 14, 2010, 11:20 p.m.	DURATION 42 minutes, 55 seconds 2 hours, 42 minutes, 59 seconds 44 minutes, 21 seconds	Training ratio: ROI size:	0.90 Ratio of the objects used for training (between 0.0 and 1.0)
IAANS Classification 9 Classification 7 Classification 5 Classification 5 Classification 1	Aug. 3, 2010, 9:34 a.m. July 26, 2010, 9:34 p.m. July 26, 2010, 5:18 p.m. July 14, 2010, 11:19 p.m. July 3, 2010, 10:41 p.m.	Scheduled July 26, 2010, 9:34 p.m. July 26, 2010, 5:19 p.m. July 14, 2010, 11:20 p.m. July 3, 2010, 10:41 p.m.	DURATION 42 minutes, 55 seconds 2 hours, 42 minutes, 59 seconds 44 minutes, 21 seconds 55 minutes, 2 seconds		0.90 Ratio of the objects used for training (between 0.0 and 1.0)

François Fleuret

The MASH project

23 / 29

VALIDATION SPECIFIC ISSUES

Performance evaluation is crucial to the MASH project, and involves the following difficulties:

- Over-fitting: The feature space is huge.
- "Meta" over-fitting: Contributors will design features specific to the problems at hand.
- Fairness: Contributions must be judged on an equal foot.
- Computational cost: (Pre-)processing can not be factorized as much as usually.
- Security: We have to compile and execute alien code on our machines.
- Intellectual property: The resulting system integrates code from multiple origins.



We can handle over-fitting with standard recipes:

- Over-fitting: regularization / feature selection methods, together with validation / cross-validation.
- "meta" over-fitting: our own data sets. Some of the samples / goal-planning tasks kept away from the contributors.

The goal-planning problems are randomized at every round (environment geometry, textures, target and avatar placement, etc.)

François Fleuret

The MASH project

25 / 29

VALIDATION FAIRNESS AND COMPUTATIONAL COST

Evaluation has to be fair to motivate contributors properly.

- We test heuristics with identical sample sets (e.g. common random seed in the basic experiments).
- We should estimate how likely a heuristic *could* have been selected during learning.
- Since the source code is public, *slight variations* should not be rewarded too much.
- Application-specific heuristics must not be penalized too much.
- Computation time has to be allocated fairly between contributors. We have setup a time-budget policy.

The platform has to provide a better analysis of the heuristics and trained predictor performance.

In particular, we are looking at methods to provide users with:

- Mistakes, sorted by severity.
- Contributions, sorted by efficiency.
- Clustering of mistakes, according to their similarities on the contributions.
- Clustering of contributions, according to their similarities on the tasks.

François Fleuret

The MASH project

27 / 29

CONCLUSION

The MASH project is a first attempt at allowing large teams to create large-scale complex learning systems.

We want to make tools to do more efficiently what we have been doing for years in ML, and to tap into a larger and more diverse population of experts.

The source code of all the feature extractors will be available over the course of the project under GPL2.

Your contribution is welcome !

THANK YOU

http://mash-project.eu

François Fleuret

The MASH project

29 / 29