

Lars Kai Hansen lkai@dtu.dk



# Machine Learning as a Service - what can be automated?

Learning from past workflows Mission of MLaaS The no free lunch theorem Decision making Representation & deep learning

Towards general workflows for key big data tasks



# Data driven services require integrated workflows!

#### DTU ML audio demos

- MIRocket (2006)
- CastSearch (2007)
  - Muzeeker(2009)
  - CoSound (2015)

#### Medical search engines

Brede Search neuroinformatics (2005) FindZebra, diagnostic queries (2013)

audio stream	speaker aution claisift Autio analysis
	Text segments

Fig. 1. The system setup. The audio stream is first processed using audio segmentation. Segments are then using an automatic speech recognition (ASR) system to produce text segments. The text is then processed using a vector representation of text and apply nonnegative matrix factorization (NMF) to find a topic space.

amata 3och Holp ode med	Sebre In Press - The search engin., *		
a dk Feddabay intarity pena			P 🛊 🦉 •
FindZebra About Mentions in press Project	8	🗭 Freedback	
indZebra In Press			
mazeora mirress			
МІТ			
Technology Review	THE	The Telegraph	
A powerful new search engine designed to help diagnose rare diseases could prove a boon for both medica and the public. – NT homoup from: the fair Jacks fairth Dyne the Outerform Single.	Danish researchers have sinveiled a free site called "FindZebra" in an effort to give doctors a new weapon to deal with tricky medical cases — The Trick Online Both" can deprive the timeses	FindZebra is designed to look through a number of selected databases of new detesses, meaning searches that could take hours if done manually take seconds. - 1% forguts the septement active dart and data is here states.	
theguardian	NewScientist	Smithsonian.com	
Most of us have had occasion to consult Dr	Frustrated patients and doctors can also turn to	(Slearch engines such as Google are not	
Google. Rather than waste a GP's time with your	FindZebra, a recently launched search engine for	designed to help a physician weed out	
embarrassing worries, just type your symptoms into a search engine. hit return and terrify	raro disoases.	possibilities behind an obscure set of symptoms [_] To fill this gap, researchers from Denmark	

	in the second	
	A 1/2 March 19 May - An Unite Cherology	+
_	-listi @ + O inpares D incores Terrine + alertite - Lanes  List	
	ErRents farade the	
_		
	and results for "arrade flat"	
	reals result for arcane free	
	Pricede Pierre	
	reprop. When, Bank Service, and a service s	
	reade Fire is an order real-band based in teatreal. Quebec which is Based round the hadrand and wife due of this buffer and Regne Charcegne. In	
	the fact fact and with about the fact of the back and recorded. The program of	
_	targe number of muck al anti-unents in addit.	
	rgeolive.unkgedet.org.intellinosi	
	Album Communition	
_	How Convertion is a Hervegian Jack Into. In 1998 the programme evended	
	How Connection is a hervegine ALL title the group very events evented be title of "young sect muticises of the year" at hervey's prestigious	
_	olds international Jazz Factoral. They have performed at the lightnessy	
	acz Pentrival, the Horth Sea Jazz Pentrival,	
	Territor State and well stream	
	wighterfreed 3 disease Duty	
	imply treating table, interactive factorial values, angleborhood (1) dream factory an alternative real, young to Canadian real.	
_	and Arcade Fire in was the third imple released from the band's debut	
	Burn, Puteral The people was released on 23 Hay. 2005 on the Rough Trade	
	scorth label. The single peoled at /26 on the US	
	Pper Coll addition for your Process	
	ware Billing	
	and Division in the second study about by Canadian type rock hand Arcade	
	ten but of the proceed could advant by Canadran more rock band would be tre. It was referred to iterate 5, 2007 in Earlings and iteration, 2007 in Herth	
	merically frough Trade and Iverge, respectively. Originally announced on	
	examples to 2004 through the basely unlights a m	

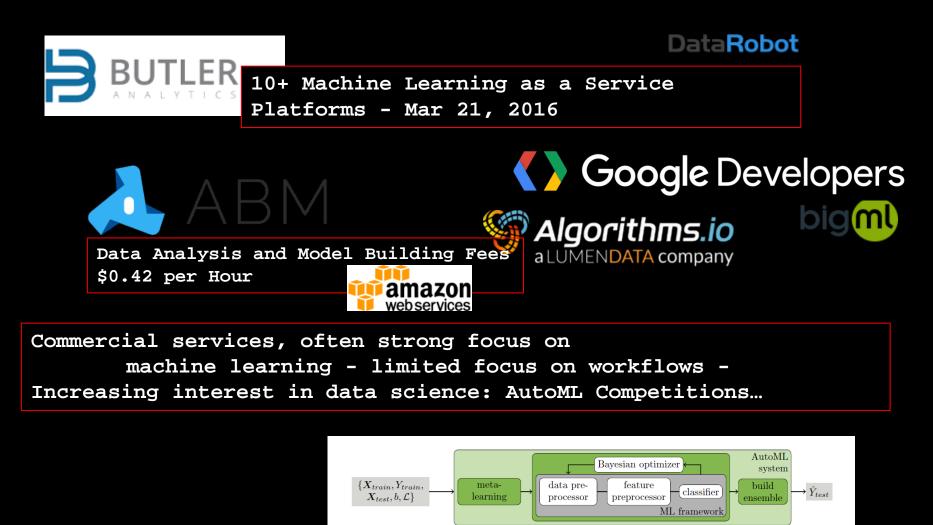
Alternative Rock         Blues         Classical         Country         Dance         Folk         Jazz         Opera & Vocal         Pop         R&B         Rap & Hip-Hop	- E					WINAMP	
Alternative Rock         Blues         Classical         Country         Dance         Folk         Jazz         Opera & Vocal         Pop         R&B         Kap & Hip-Hop	Eile	Play	Options	View	Help		
Blues       Classical       Country       Dance       Folk       Jazz       Opera & Vocal       Pop       R&B       Kap & Hip-Hop	(F	Detach V	'isualizer				
Classical Country Dance Folk Jazz Opera & Vocal Pop R&B R&B						Alternative Rock	
Country Dance Folk Jazz Opera & Vocal Pop R&B R&B Rap & Hip-Hop						Blues	
Dance       Folk       Jazz       Opera & Vocal       Pop       R&B       Rap & Hip-Hop					]	Classical	
Folk       Jazz       Opera & Vocal       Pop       R&B       Rap & Hip-Hop						Country	
Jazz           Opera & Vocal           Pop           R&B           Rap & Hip-Hop						Dance	
Opera & Vocal           Pop           R&B           Rap & Hip-Hop						Folk	
Pop           R&B           Rap & Hip-Hop						Jazz	
R&B           Rame           Ram           Ram <th></th> <th></th> <th></th> <th></th> <th></th> <th>Opera &amp; Vocal</th> <th></th>						Opera & Vocal	
Rap & Hip-Hop						Рор	
						🕺 🕺 🕹 🕹 🕹 🕹 🕹 🕹	
Rock						Rock	

B:88

MUDDY WATERS



## Machine Learning as a Service



Feurer M, Klein A, Eggensperger K, Springenberg J, Blum M, Hutter F. Efficient and robust automated machine learning. In Advances in Neural Information Processing Systems 2015 pp. 2962-2970.)



#### The mission of MLaaS

Basic machine learning functions are well defined, have wellunderstood representations and learning curves. Large number of open source tools are available to the data scientist

Structure, "unsupervised learning" p(measurement|structure)
exploratory analysis and simulation, novelty detection, cleaning

Tuning complexity to given application still a challenge... MLaaS: Help data scientist to tune, validate, understand ML





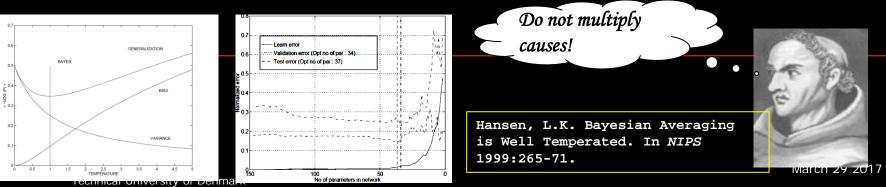
## No free lunch in data science

**Bad news:** Without modeling "a priori" you can learn nothing! Without priors, you are left with "uniform inference" (Wolpert, D.H., 1996. The lack of a priori distinctions between learning algorithms. Neural computation, 8(7), pp.1341-1390.)

GOOD NEWS: Well-understood means for optimizing models for relevant & "physically" relevant priors/representations - typical tuning complexity

Bayes vs Cross-validation - Bayes is optimal in the rare case when you got the prior right (no complexity tuning....

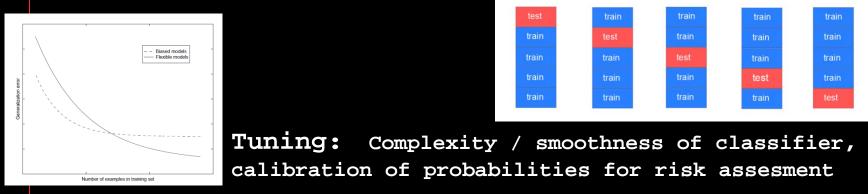
Cross-validation optimizes when prior is uncertain.... Often used forms of prior information used: Smoothness in feature space, sparsity / dimensional, transfer learning...





Classification, signal detection, discrimination

- Generalization: Predcition errors on test
- Training, validation, test simulated by resampling



# Potential for automatic tuning: High, good examples provided by Hutter group:

"In this work we introduce a robust new AutoML system based on scikit-learn (using 15 classifiers, 14 feature preprocessing methods, and 4 data preprocessing methods, giving rise to a structured hypothesis space with 110 hyperparameters). This system, which we dub auto-sklearn, improves on existing AutoML methods by automatically taking into account past performance on similar datasets, and by constructing ensembles from the models evaluated during the optimization"

Feurer M, Klein A, Eggensperger K, Springenberg J, Blum M, Hutter F. Efficient and robust automated machine learning. In Advances in Neural Information Processing Systems 2015 pp. 2962-2970.)

0.02

#### Gaussian (and other) mixtures

(Hinton, Geoffrey E., and Ruslan R. Salakhutdinov. "Reducing the dimensionality of data with neural networks." Science 313.5786 (2006): 504-507.)

Pretraining for neural networks: Hinton's revelation

Tuning: Complexity of structure (e.g. #clusters, latent variable dimensionality,...)

Potential for automatic tuning: High, work is ongoing (F. Zdyb)



Generative adverserial networks

Non-parametrics / kernel methods



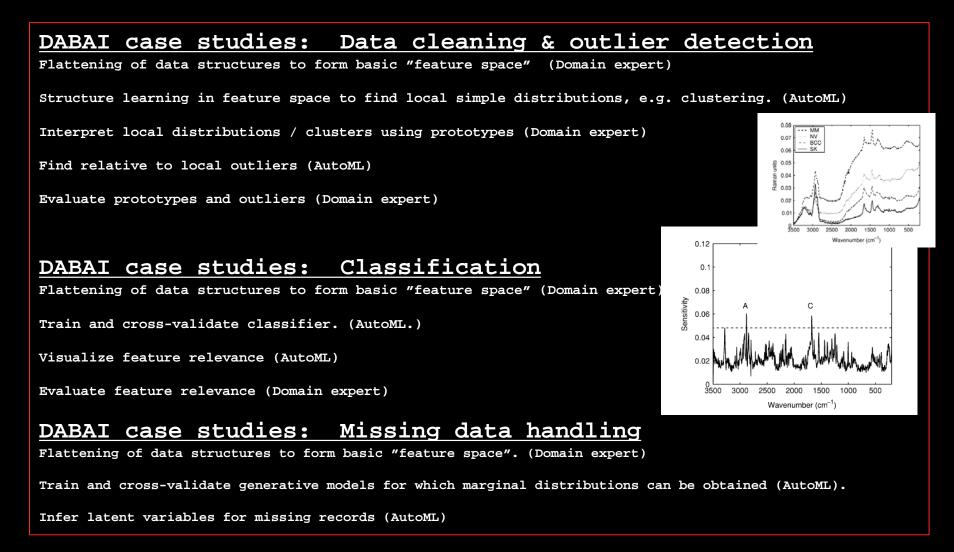


representation learning p( measurement | structure )

### Structure &

Typical use cases:

## Workflows for prototypical applications



Sigurdsson, S., Philipsen, P.A., Hansen, L.K., Larsen, J., Gniadecka, M. and Wulf, H.C., 2004. Detection of skin cancer by classification of Raman spectra. *IEEE transactions on biomedical engineering*, 51(10):1784-1793.

Data science with machine learning as a service



Data driven services require integrated workflows!

Multiple steps ... some interactive, some automated

Interactive: Data preparation / exploration

Automate: Machine learning representations

Automate: Machine learning decision support systems

Interactive: Model interpretation / decision making

and New York	a tark Times Economist.com Newsweett
	willie Friends Inbox
-	What's on your mind?
Energy .	
Elmonton, M	
	Add: 🕤 Link 💽 Photos 😒 Videos 🜰 Same