

Towards interactive Machine Learning for solving complex problems

Andreas Holzinger

Holzinger Group, Research Unit HCI-KDD, Institute for Medical Informatics,
Statistics and Documentation, Medical University Graz

&

Institute of Information Systems and Computer Media,
Graz University of Technology

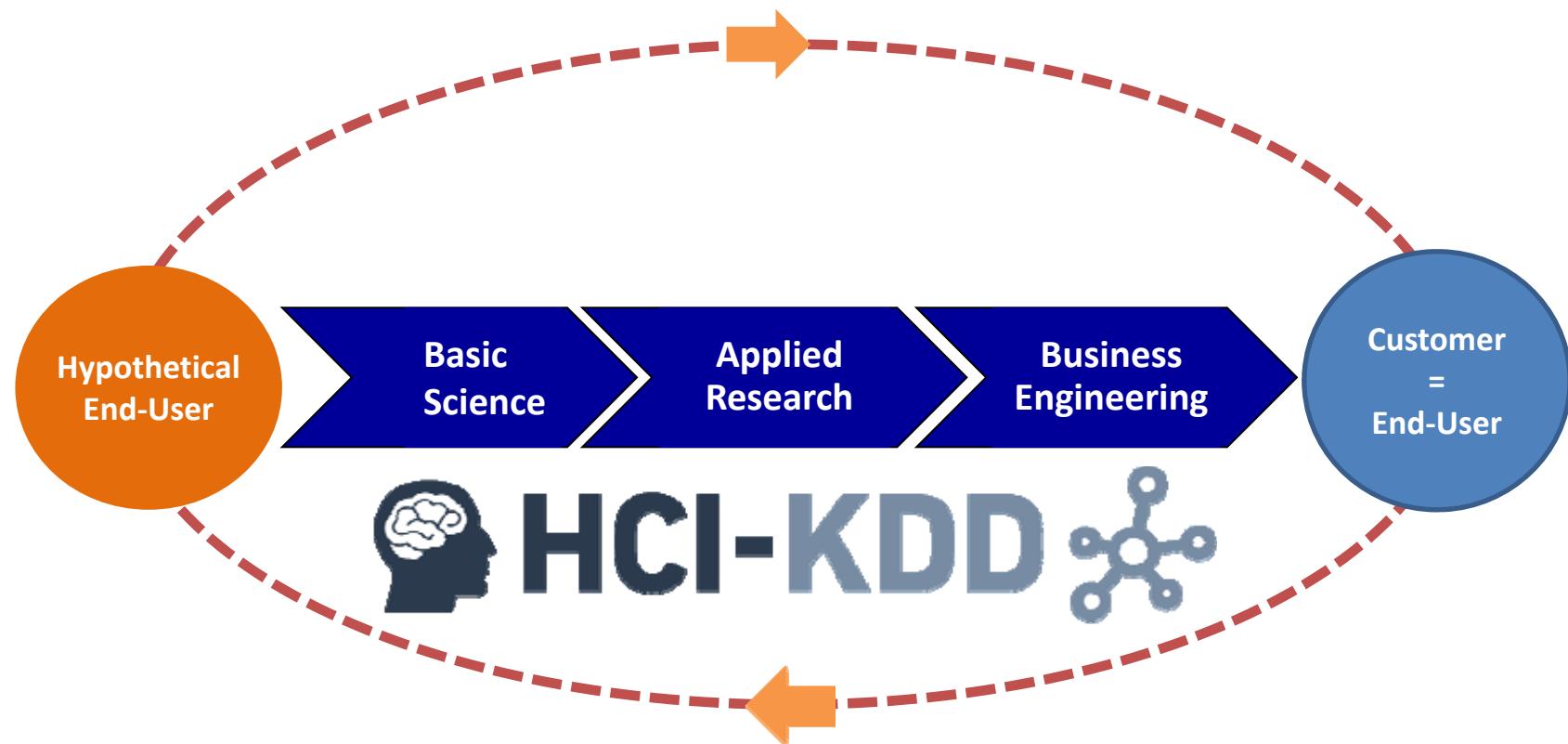


- The HCI-KDD approach
- ML and Health Informatics
- ML state-of-the-art
- aML versus iML
- A few examples of iML
- Future Outlook



a.holzinger@hci-kdd.org

Science is testing crazy ideas – Engineering is putting these ideas into Business



Holzinger, A. 2011. Successful Management of Research and Development, Norderstedt: BoD.

Interactive

Data Mining



⑥ Data Visualization

② Learning Algorithms

Knowledge Discovery

① Data Mapping

Prepro-
cessing

Data
Fusion



GDM ③ *Graph-based Data Mining*

TDM ④ *Topological Data Mining*

EDM ⑤ *Entropy-based Data Mining*

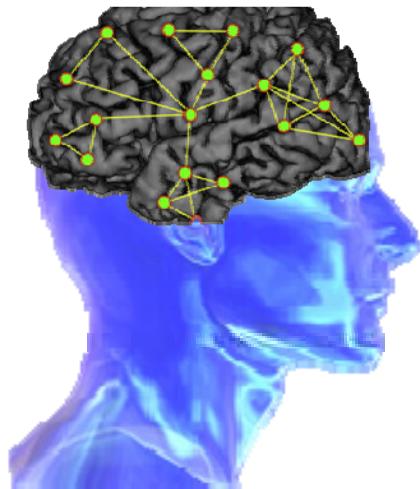
Privacy, Data Protection, Safety and Security

© a.holzinger@hci-kdd.org

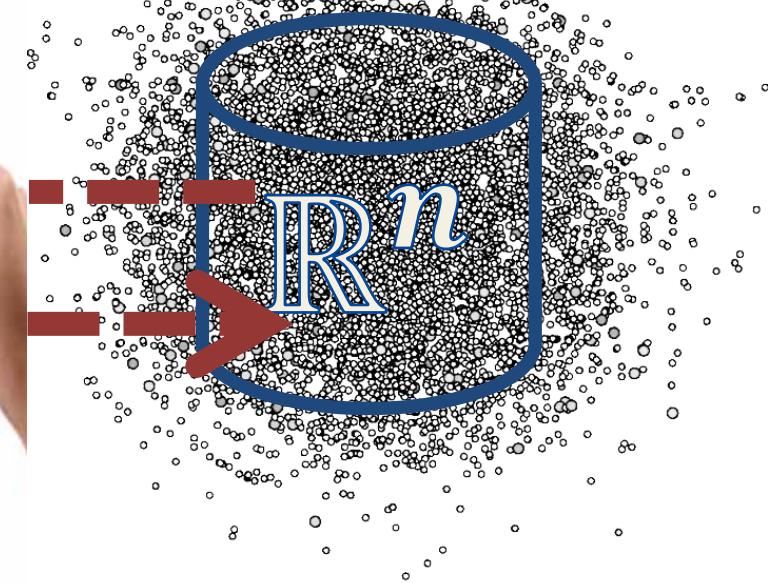
7

Holzinger, A. 2014. Trends in Interactive Knowledge Discovery for Personalized Medicine: Cognitive Science meets Machine Learning. IEEE Intelligent Informatics Bulletin, 15, (1), 6-14.

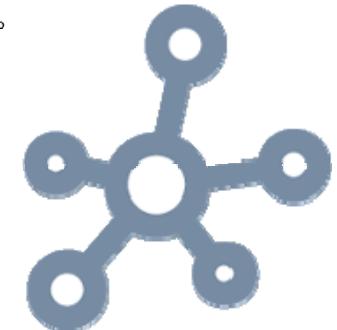
Human intelligence (Cognitive Science)



Machine intelligence (Computer Science)



HCI-KDD



Holzinger, A. (2013). Human–Computer Interaction & Knowledge Discovery (HCI-KDD): What is the benefit of bringing those two fields to work together? In: Lecture Notes in Computer Science 8127 (pp. 319-328)



- CS aims to reverse engineer **human intelligence**;
- ML provides powerful sources of insight into **how machine intelligence** is possible.
- CS therefore raises challenges for, and draws inspiration from ML;
- Insights about the human mind may help inspire **new directions for ML** ...

A photograph of a surgical team in an operating room. Several surgeons wearing blue scrubs and caps are gathered around a patient, performing a laparoscopic procedure. One surgeon's hands are visible, holding a surgical instrument. Above them, four large, illuminated surgical lights are mounted on a track. Two monitors are positioned on stands, displaying close-up video feeds of the surgery from different angles. The background shows shelves stocked with medical supplies.

Application Area: Health Informatics

Why is this application area complex ?



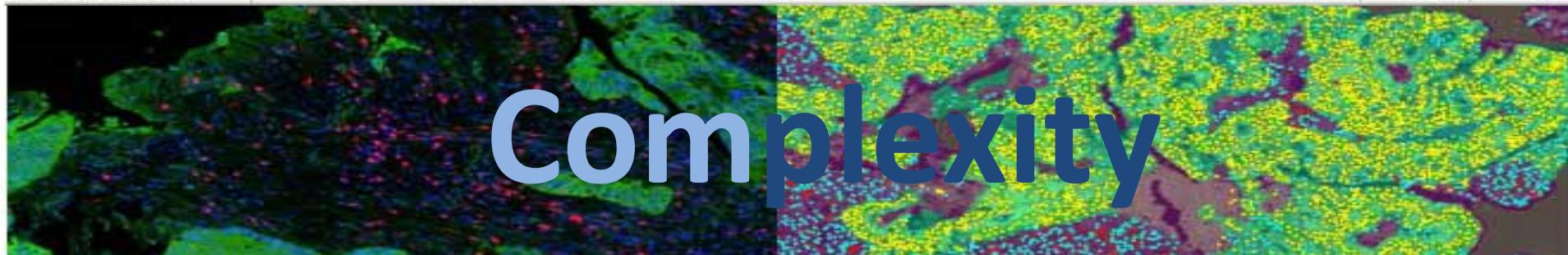
Our central hypothesis: Information may bridge this gap

Holzinger, A. & Simonic, K.-M. (eds.) 2011. *Information Quality in e-Health. Lecture Notes in Computer Science LNCS 7058*, Heidelberg, Berlin, New York: Springer.



Where is the
problem in building
this bridge?

		μ_{12}	T^*	μ_{13}		μ_{14}		μ_{15}		μ_{16}		μ_{17}		μ_{18}		μ_{19}		μ_{20}	
MARHY0478	299	YsgDGWRIGGSIEQQNWESELEDEFsgdsik~~d	qsvaSGNRIGFDD	LIPPE	GAEYQLNk	NFAVRGGVA													
FadL	300	AFWTVVLAATW	QGQD	acsL	-scit	EKHGF	D	YRIL	LG	TY	WT	NT							RTGIA
IbuX	312	Fn~DQLSVSADYQRVFK	N	MKDmnvgivqsgsaanld	SLPQNYRD	ISVFG	G	AEYRYNaK	W	IFRGGFH									
TodX	309	Fn~ERWVVAADIKRAYWDVMDSmnvafis~~qlggid	VALPHRYQDITV	VASIGTAYKYNn	DLTL	RAGYS													
Total pos/pS		16	16	5	21	21	21	5	26	26	26	26	5	31					
Total Infusionen		8	116	8	125	125	125	125	42	166	166	17	183	8	191	191	17	17	
Total Meds (pos+iv)			4		4		4		4		6		6		6				
Total Perfusoren		1	9	1	10	10	10	10	5	15	15	2	17	1	18	18	2	2	
Total Meds+Perfusoren		1	13	1	14		14	14	7	21	21	2	23	1	24	24	2	2	
Total Blut																			
Total Harn			43		43		83		4		12		12		34				
Harmenge/Zeit									1/4						1/4				
Harn/kg/S									2/6						0				
Total Ma-Dosis			6		6		6		6		6		6		6				
Total Blut																			
Total Ein		9	145	9	154	5	159	159	159	54	213	213	19	232	9	241	5	246	18
Total Aus		49	49	40	89	89	89	89	29	118	118	118	22	140	140				
Nettobilanz 24h		+96	+105	+70	+70	+70	+70	+70	+95	+95	+114	+114	+101	+106	+106	+18			



Uncertainty

Probabilistic Information $p(x)$

Of great help was ...

the following list of resources that were of great help during the development of the system. The list is ordered by relevance to the system development process.

- **System Requirements**:
The system requirements were defined in a document titled "System Requirements Document" (SRD). The SRD was developed by the project team and reviewed by the stakeholders. It contained a detailed description of the system's requirements, including functional and non-functional requirements, and was used as a reference throughout the development process.
- **Design Patterns**:
The design patterns were used to guide the system architecture and design. The patterns were selected based on their relevance to the system requirements and the project context. They provided a set of best practices for system design, which helped to ensure the system's quality and maintainability.
- **Code Review**:
The code review process was used to identify and fix errors and bugs in the system code. The code review was conducted by a team of experienced developers who reviewed the code and provided feedback to the developer who wrote it. This process helped to ensure the system's quality and reliability.
- **Testing**:
The testing process was used to verify the system's functionality and performance. The testing was conducted by a team of experienced testers who tested the system against the system requirements. They identified and fixed errors and bugs in the system code, and provided feedback to the developer who wrote it. This process helped to ensure the system's quality and reliability.
- **Deployment**:
The deployment process was used to move the system from the development environment to the production environment. The deployment was conducted by a team of experienced system administrators who managed the deployment process. They ensured that the system was deployed correctly and without any issues.
- **Monitoring**:
The monitoring process was used to track the system's performance and identify any issues or problems. The monitoring was conducted by a team of experienced system administrators who monitored the system's performance and provided feedback to the developer who wrote it. This process helped to ensure the system's quality and reliability.
- **Documentation**:
The documentation process was used to create documentation for the system. The documentation was created by a team of experienced writers who documented the system's architecture, design, and implementation. They provided a set of guidelines for system users, which helped to ensure the system's quality and maintainability.
- **Feedback**:
The feedback process was used to collect feedback from the system users. The feedback was collected through various channels, such as surveys, user interviews, and user feedback forms. The feedback was used to identify any issues or problems with the system and to make improvements to the system.
- **Iteration**:
The iteration process was used to refine the system's requirements and design. The iteration was conducted by a team of experienced developers who iterated on the system requirements and design. They identified and fixed errors and bugs in the system code, and provided feedback to the developer who wrote it. This process helped to ensure the system's quality and maintainability.

Probabilistic Information p(x)



Bayes, T. (1763). An Essay towards solving a Problem in the Doctrine of Chances (Postum communicated by Richard Price). Philosophical Transactions, 53, 370-418.

$$p(x_i) = \sum P(x_i, y_j)$$

Thomas Bayes
1701 - 1761

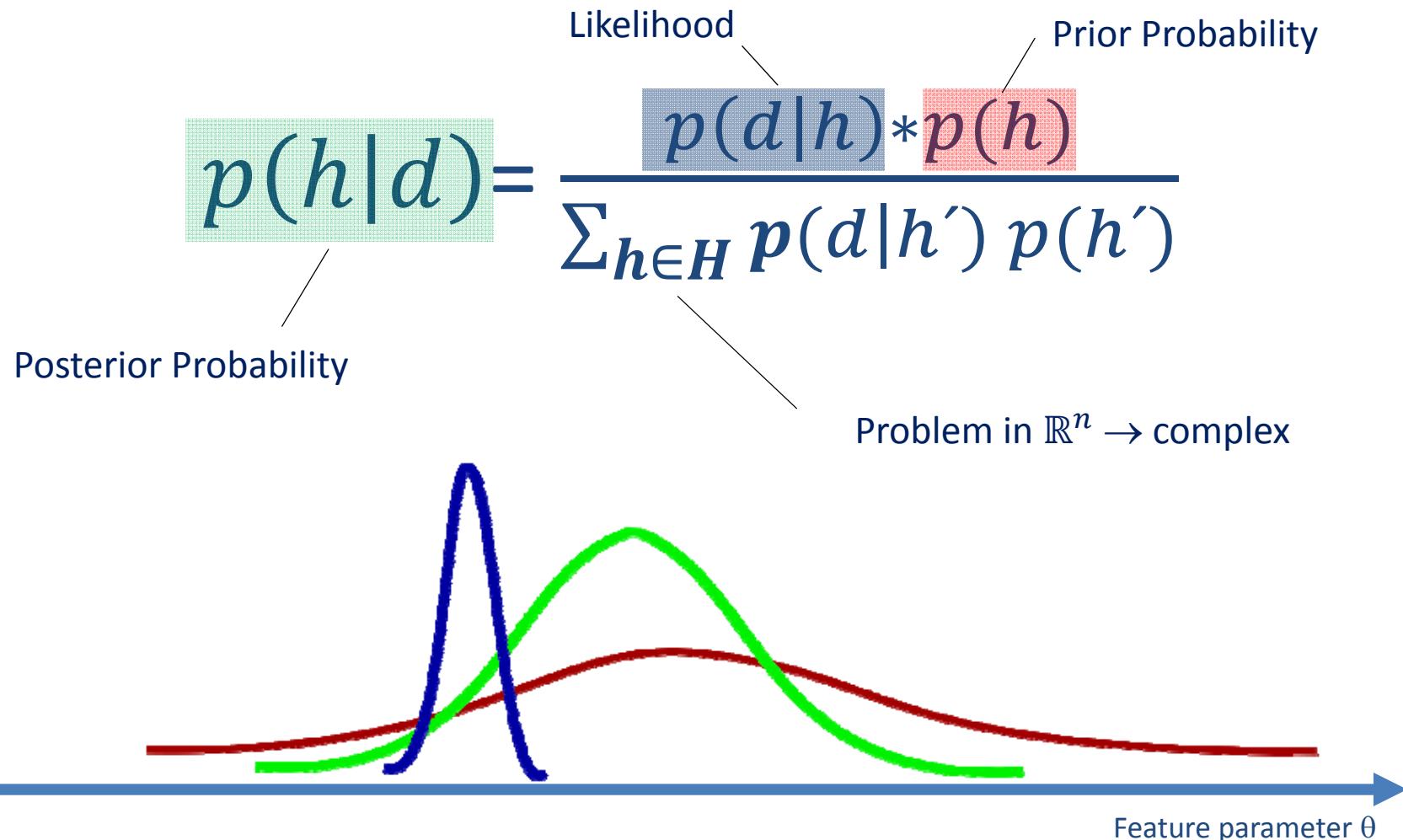
$$p(x_i, y_j) = p(y_j|x_i)P(x_i)$$

Bayes' Rule is a corollary of the Sum Rule and Product Rule:

$$p(x_i|y_j) = \frac{p(y_j|x_i)p(x_i)}{\sum p(x_i, y_j)p(x_i)}$$

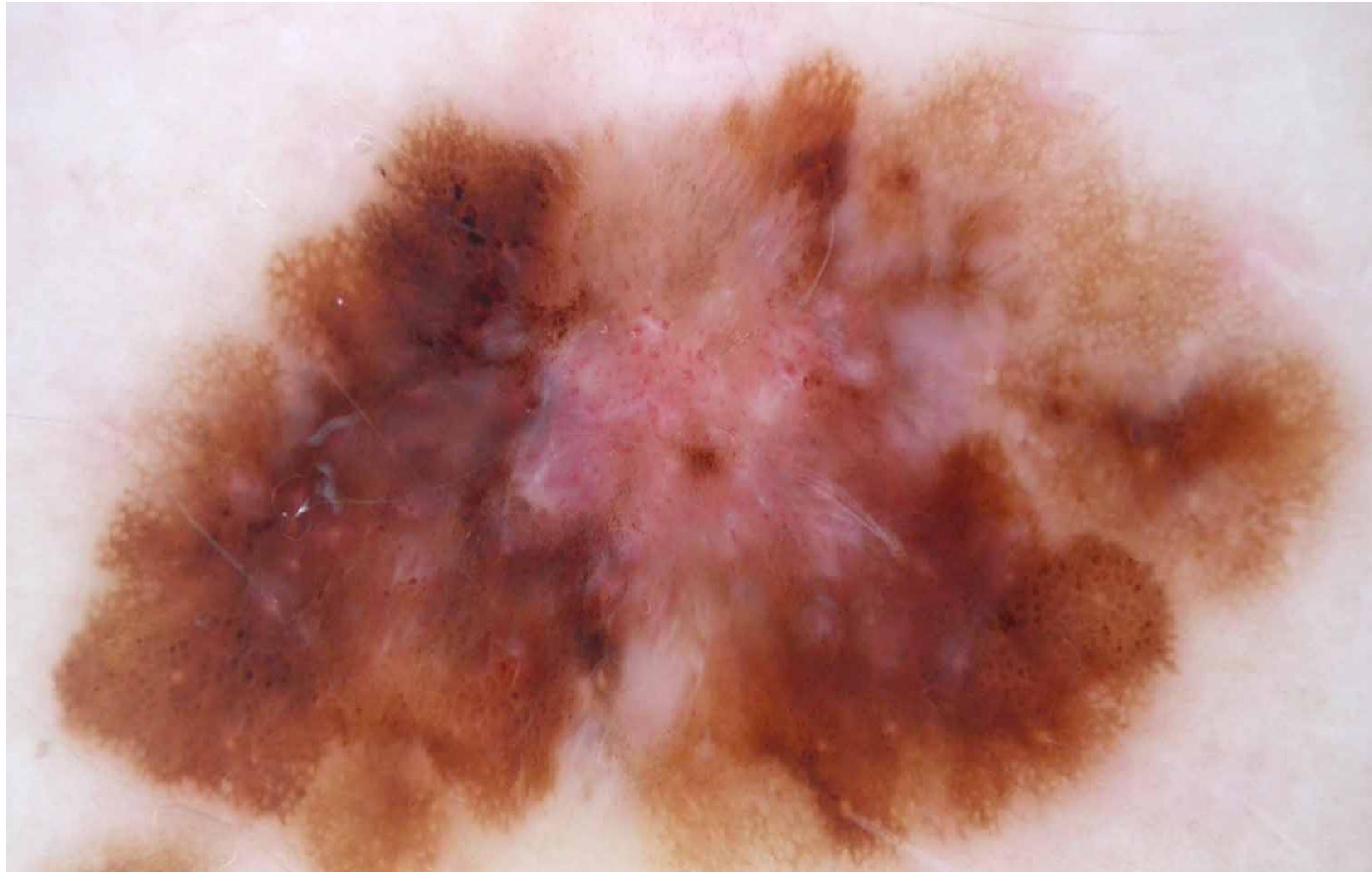
Barnard, G. A., & Bayes, T. (1958). Studies in the history of probability and statistics: IX. Thomas Bayes's essay towards solving a problem in the doctrine of chances. Biometrika, 45(3/4), 293-315.

$d \dots \text{data}$ $h \dots \text{hypothesis}$ $H \dots \{H_1, H_2, \dots, H_n\}$ $\forall h, d \dots$



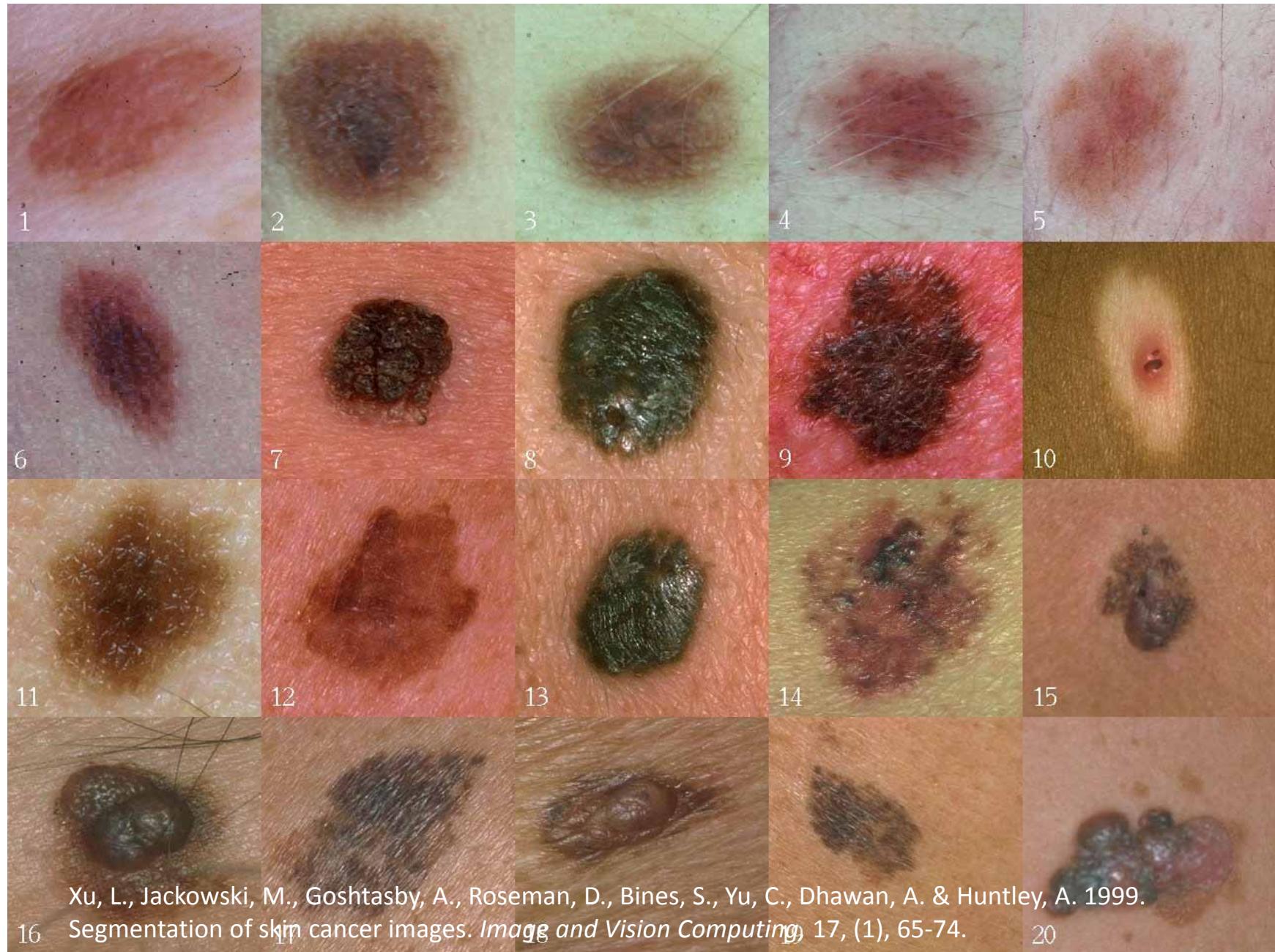
Biomedical Example

To find structural anomalies in such data ...



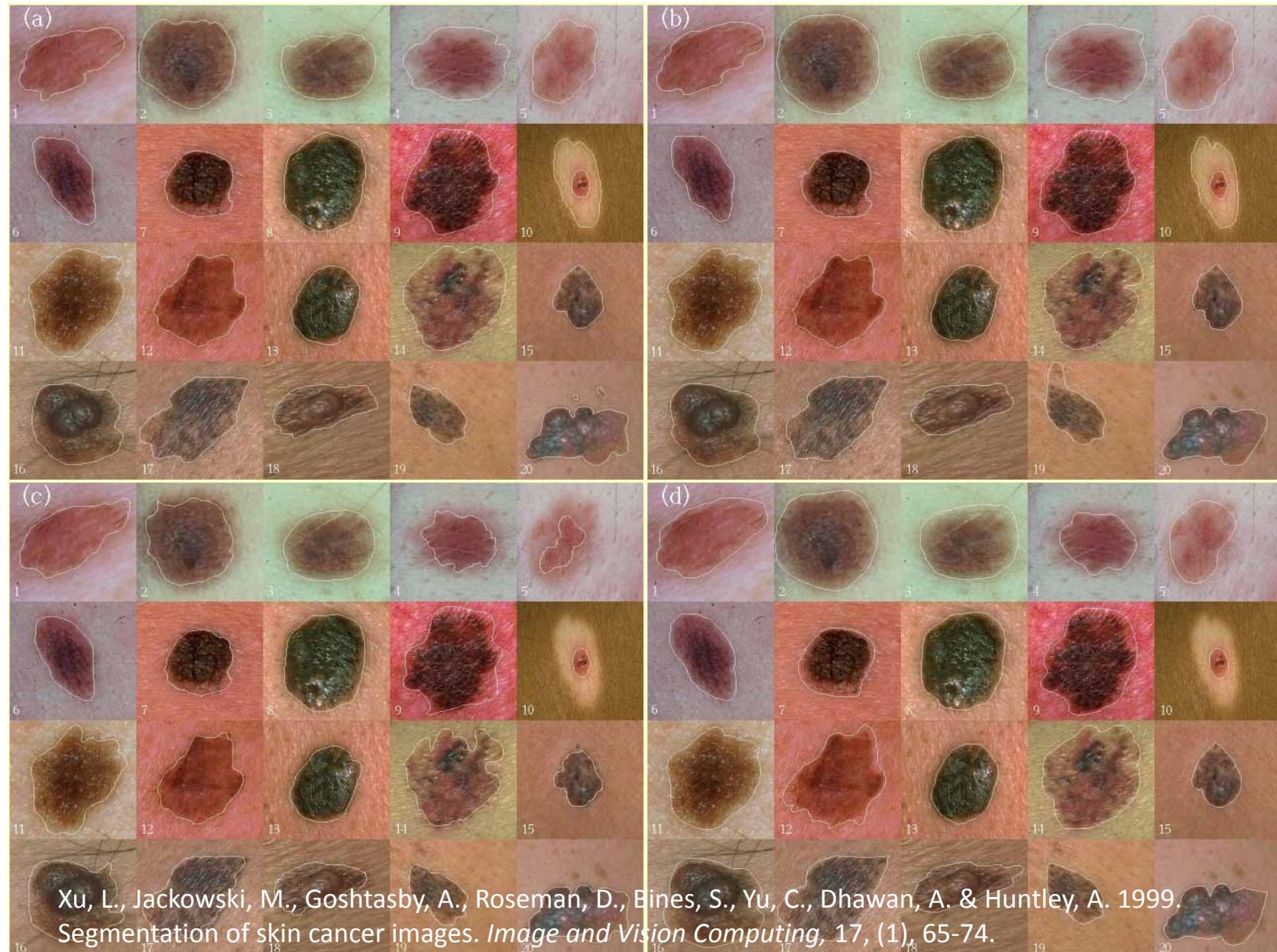
Holzinger, A., Malle, B., Bloice, M., Wiltgen, M., Ferri, M., Stanganelli, I. & Hofmann-Wellenhof, R. 2014. On the Generation of Point Cloud Data Sets: Step One in the Knowledge Discovery Process. In: *Lecture Notes in Computer Science, LNCS 8401*. Berlin Heidelberg: Springer, pp. 57-80, doi:10.1007/978-3-662-43968-5_4.

Learning from previous examples ...



Xu, L., Jackowski, M., Goshtasby, A., Roseman, D., Bines, S., Yu, C., Dhawan, A. & Huntley, A. 1999.
16 Segmentation of skin cancer images. *Image and Vision Computing*, 17, (1), 65-74.

The more examples we have the better ...



Big Data is good for automatic Machine Learning

$$\mathcal{D} = x_{1:n} = \{x_1, x_2, \dots, x_n\} \quad p(\mathcal{D}|\theta)$$

$$p(\theta|\mathcal{D}) = \frac{p(\mathcal{D}|\theta) * p(\theta)}{p(\mathcal{D})}$$

$$posterior = \frac{likelihood * prior}{evidence}$$

The inverse probability allows to learn from data, infer unknowns, and make predictions

$$\max_{\mathbf{x} \in \mathcal{A} \subset \mathbb{R}^d} f(\mathbf{x})$$

$$p(h|d) \propto p(\mathcal{D}|\theta) * p(h)$$

$$p(f(x)|\mathcal{D}) \propto p(\mathcal{D}|f(x)) * p(f(x))$$

- Machine Learning is the development of algorithms which can **learn from data**
- assessment of **uncertainty**, making **predictions**
- **Automating automation** - getting computers to **program themselves** – let the data do the work!



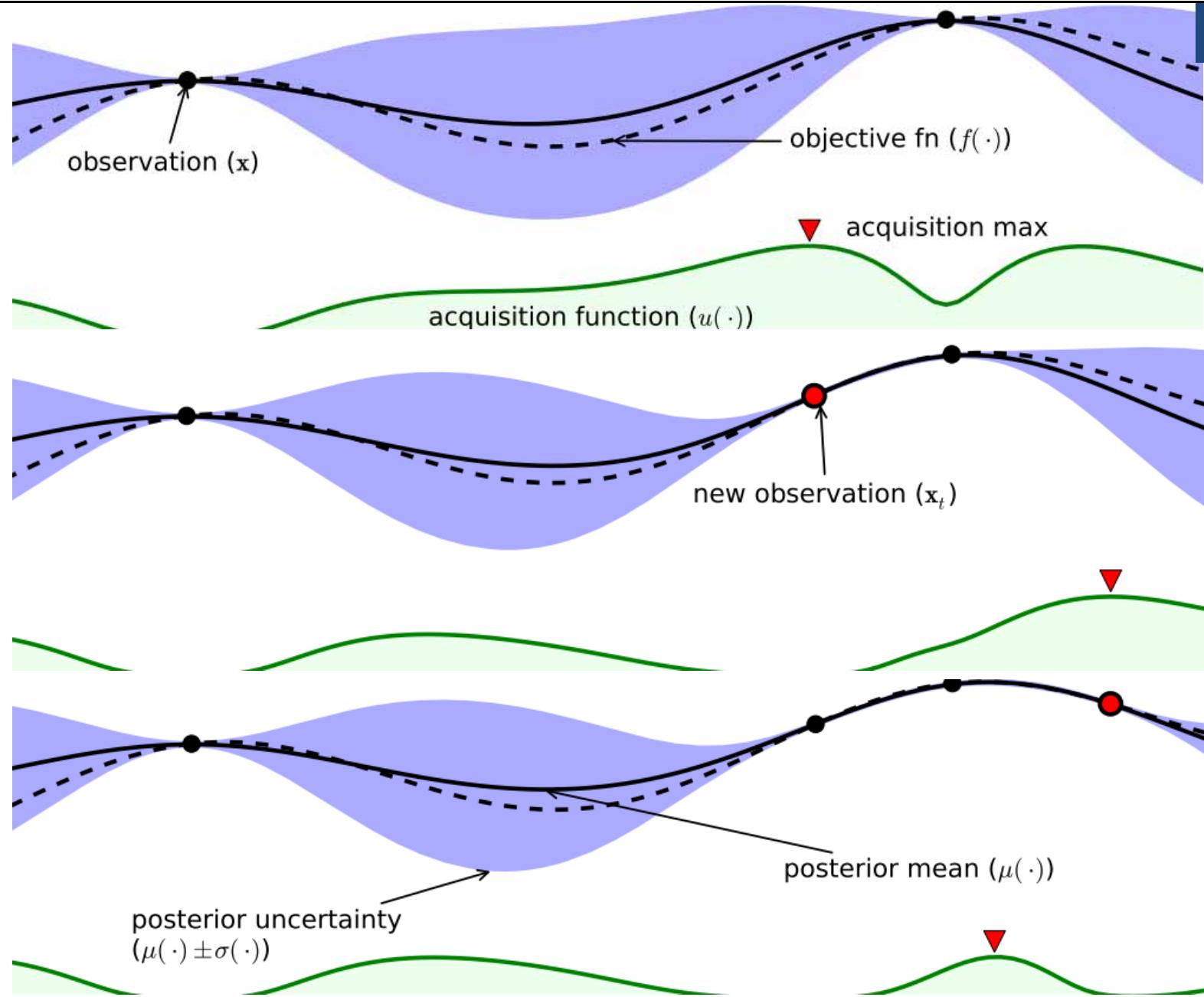
- **Newton, Leibniz, ... developed calculus – mathematical language for describing and dealing with rates of change**
- **Bayes, Laplace, ... developed probability theory - the mathematical language for describing and dealing with uncertainty**



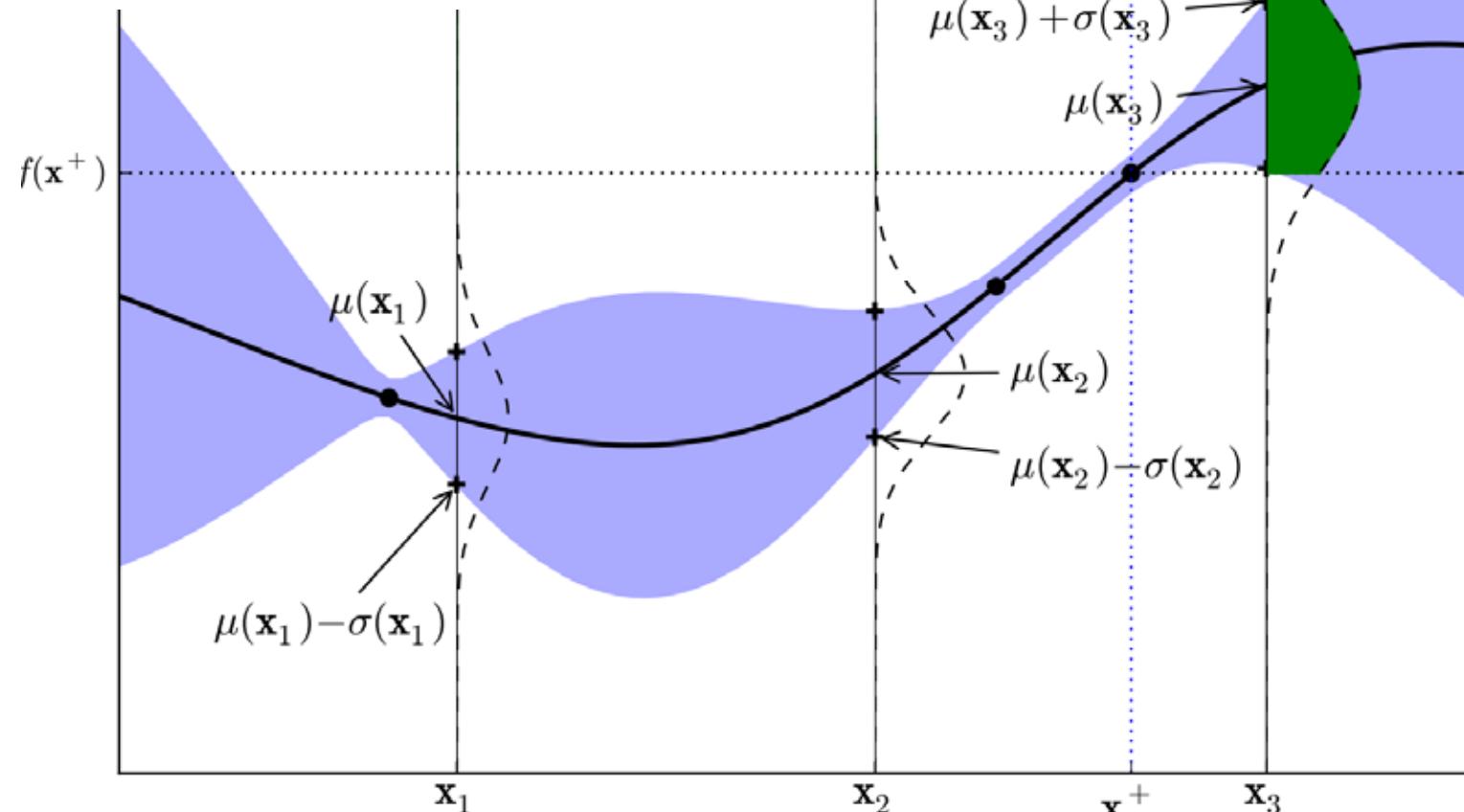
Gaussian Processes

From Bayesian Optimization to Gaussian Process (GP) approximation

Brochu, E., Cora, V. M. & De Freitas, N. 2010. A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning arXiv:1012.2599.

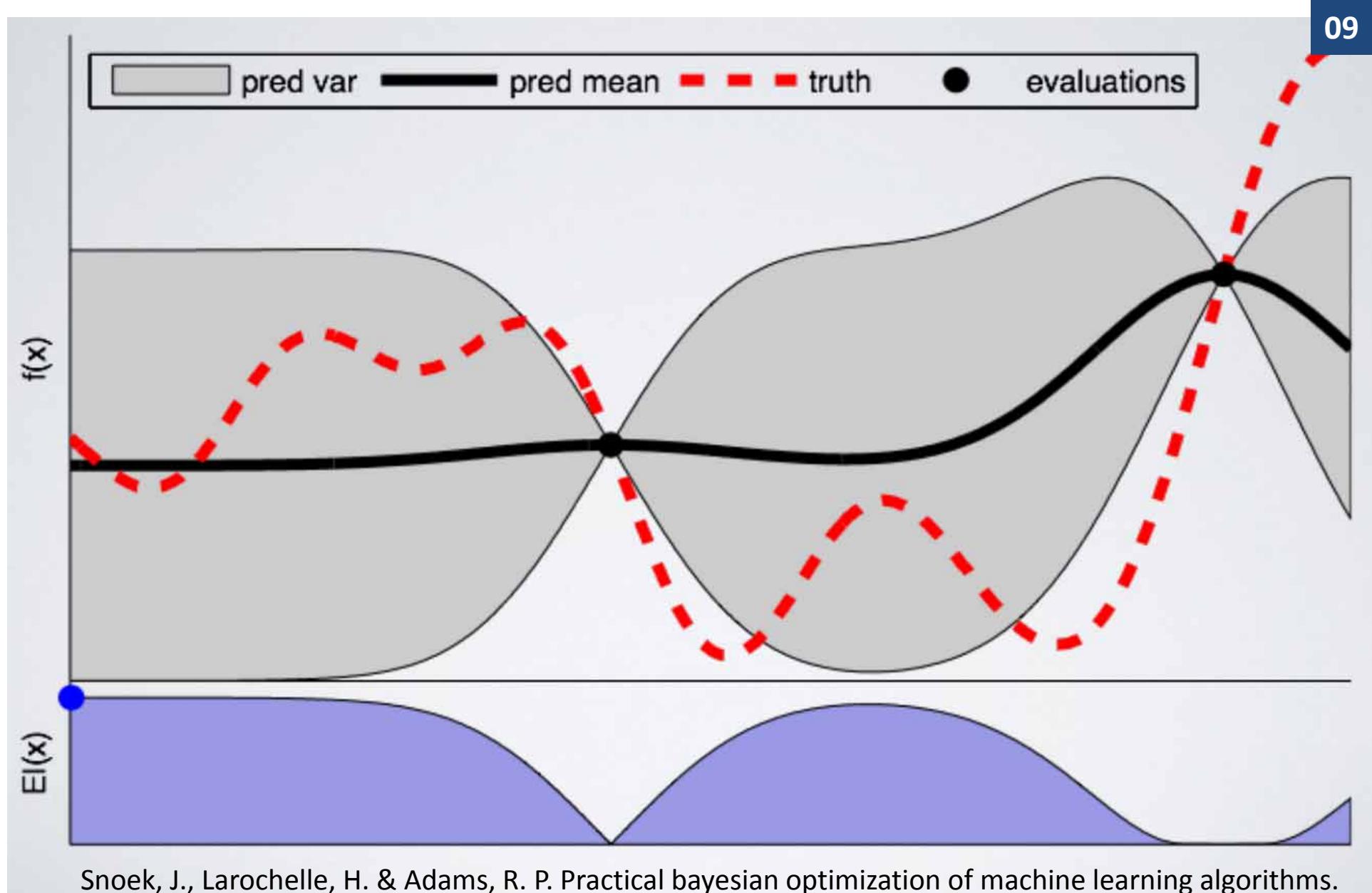


$$\text{GP posterior} \quad p(f(x)|\mathcal{D}) \propto \text{Likelihood} \underbrace{p(\mathcal{D}|f(x))}_{\text{GP prior}} p(f(x))$$

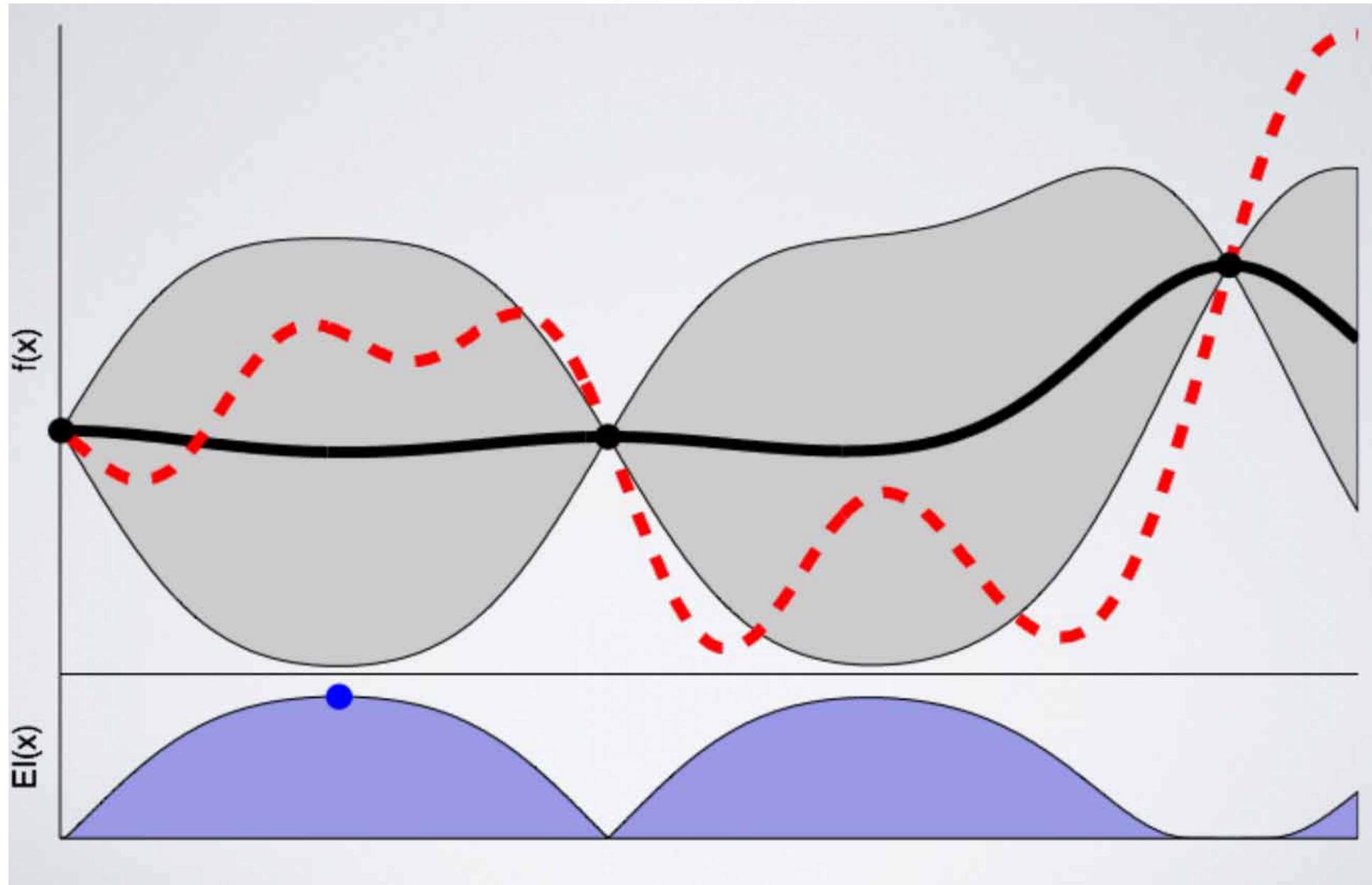


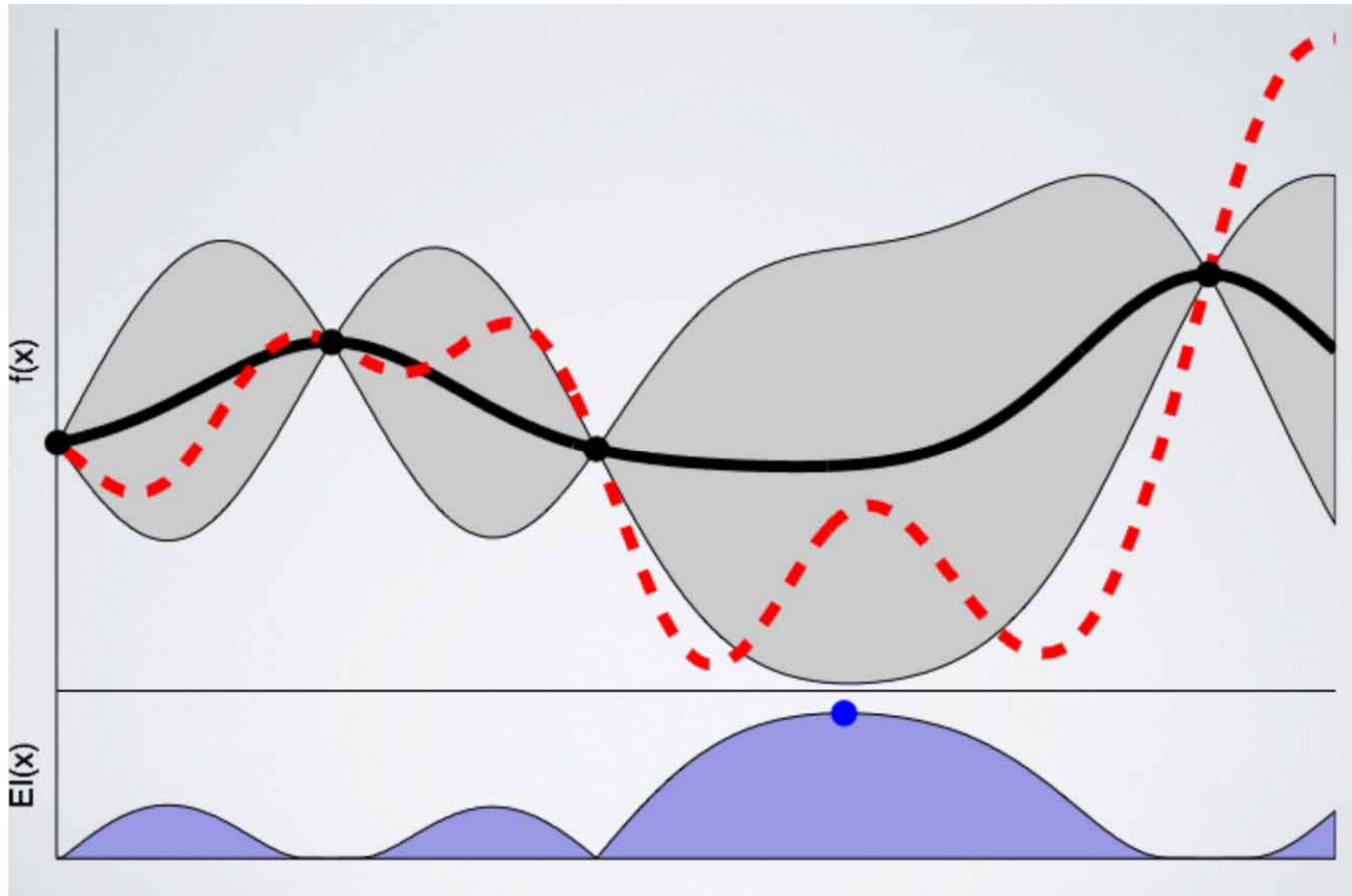
Brochu, E., Cora, V. M. & De Freitas, N. 2010. A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. arXiv:1012.2599.

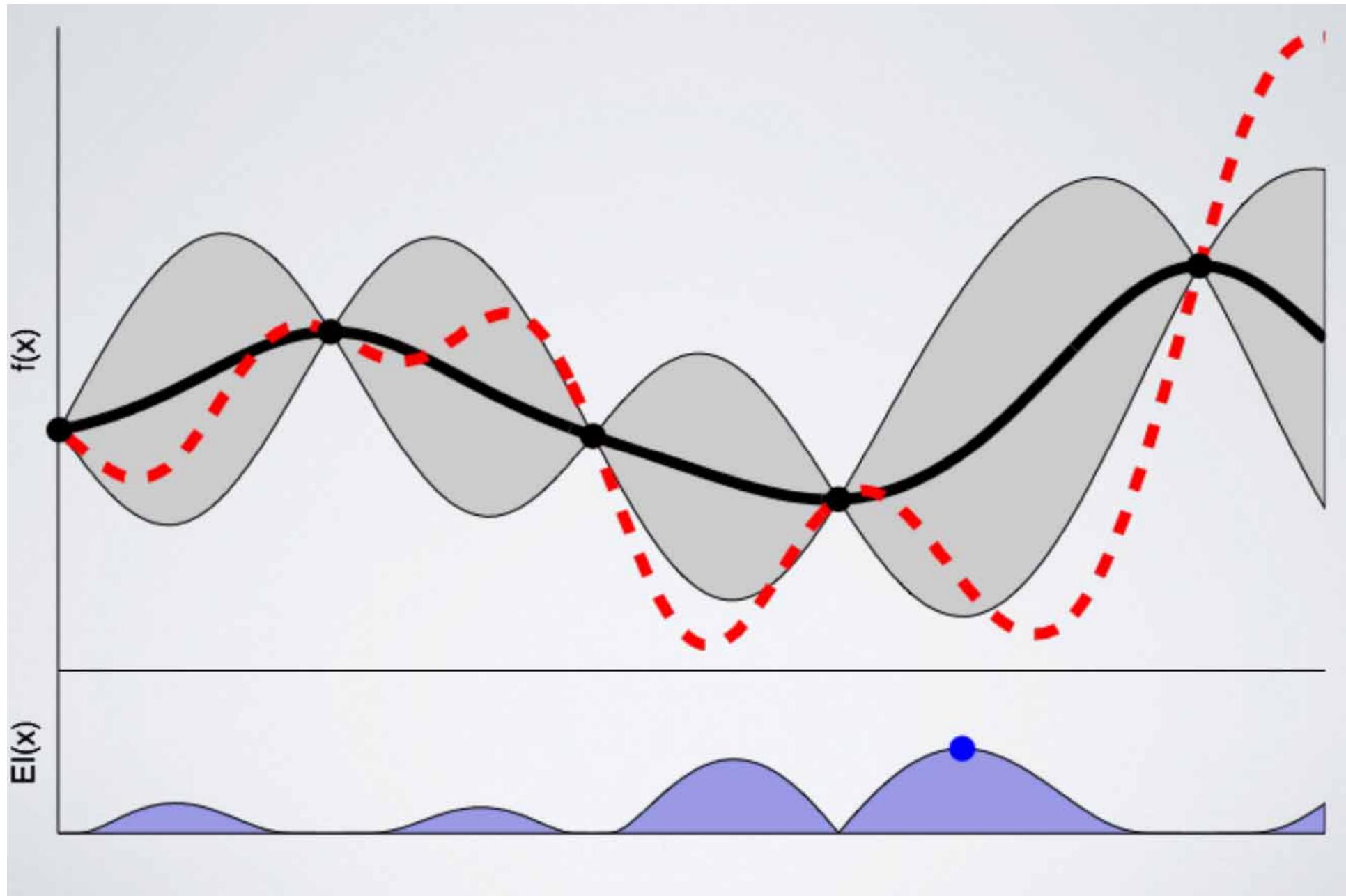
Demo on how Bayesian Optimization works ...

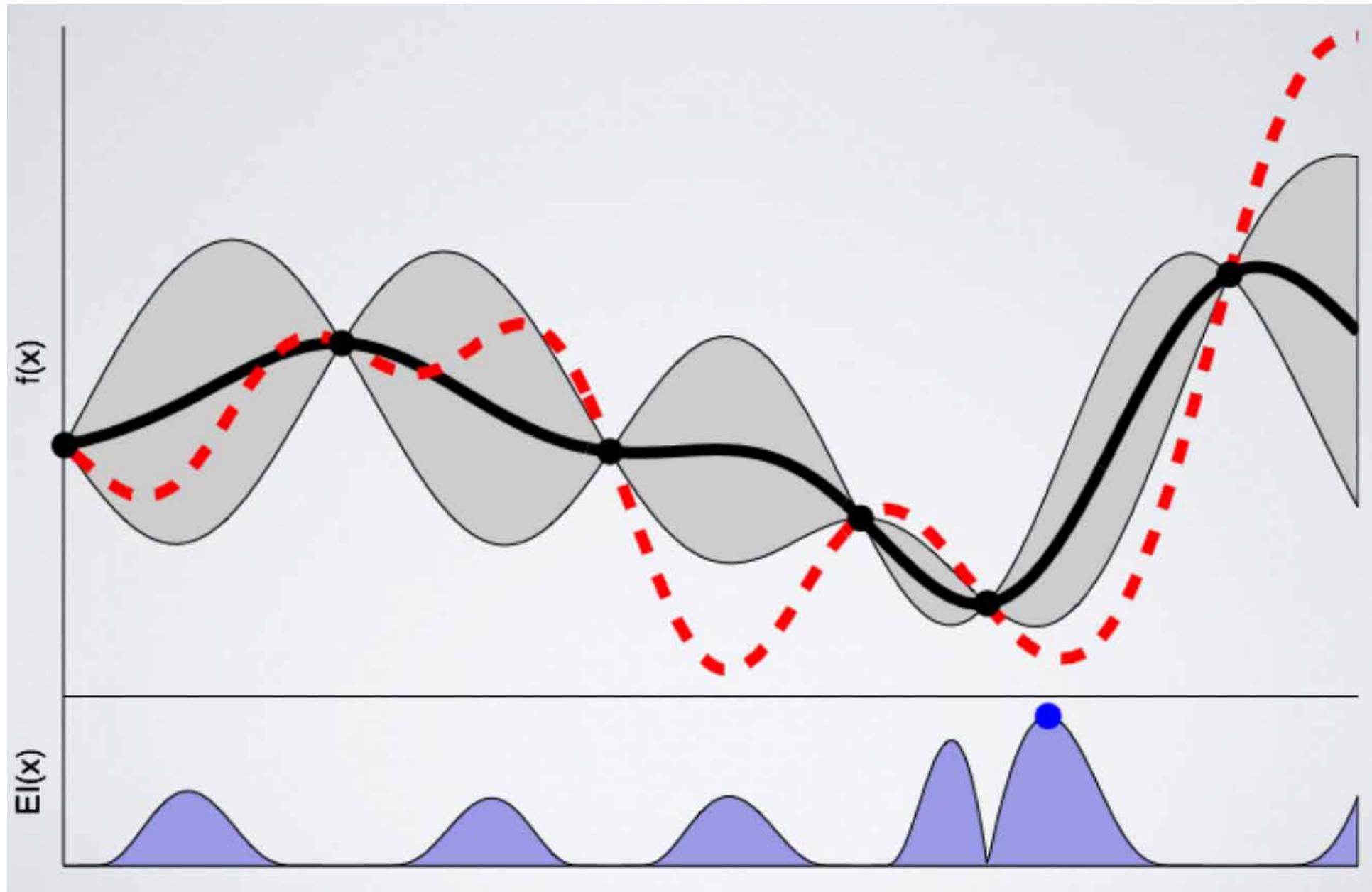


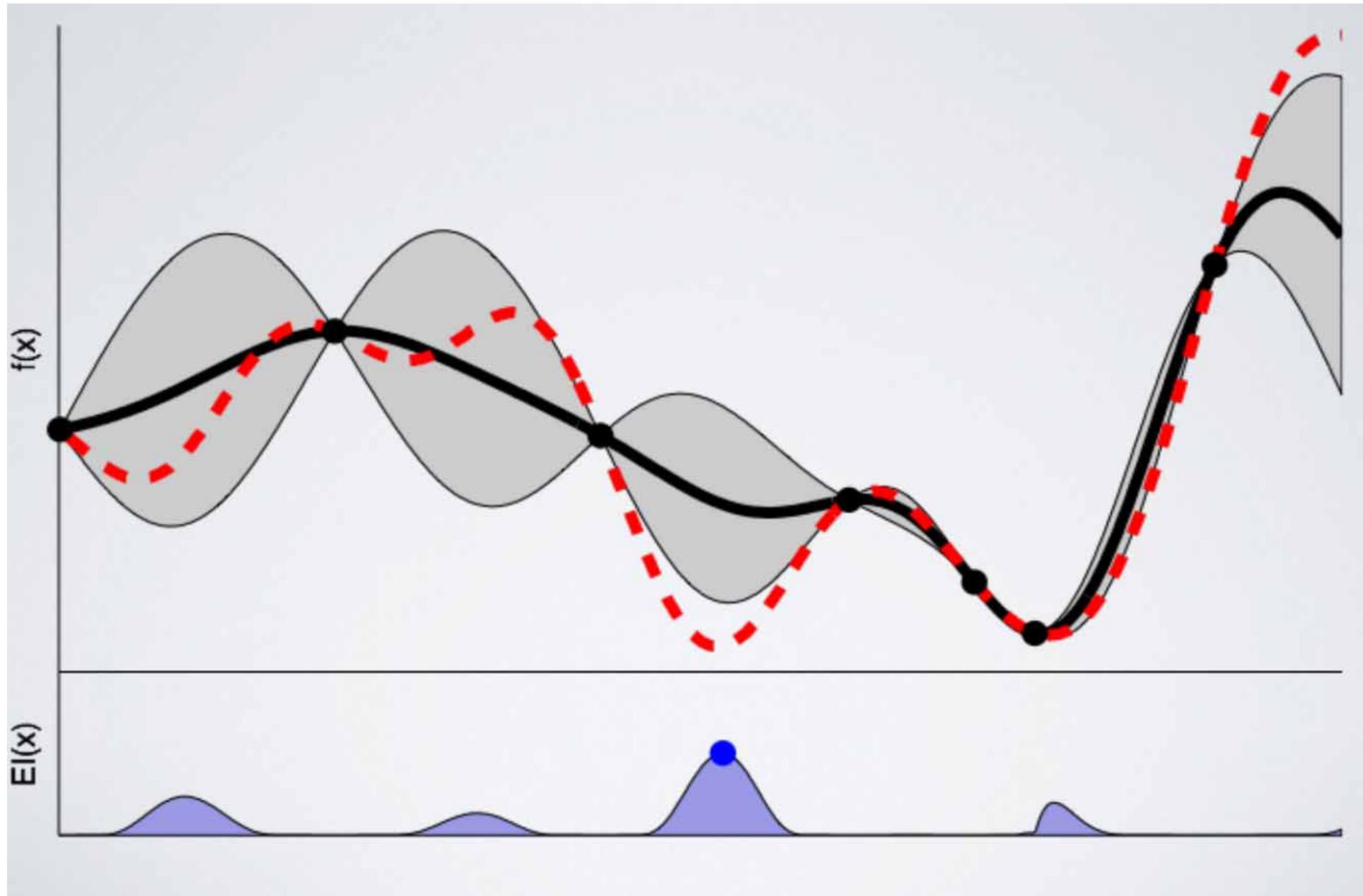
Snoek, J., Larochelle, H. & Adams, R. P. Practical bayesian optimization of machine learning algorithms.
Advances in neural information processing systems, 2012. 2951-2959.

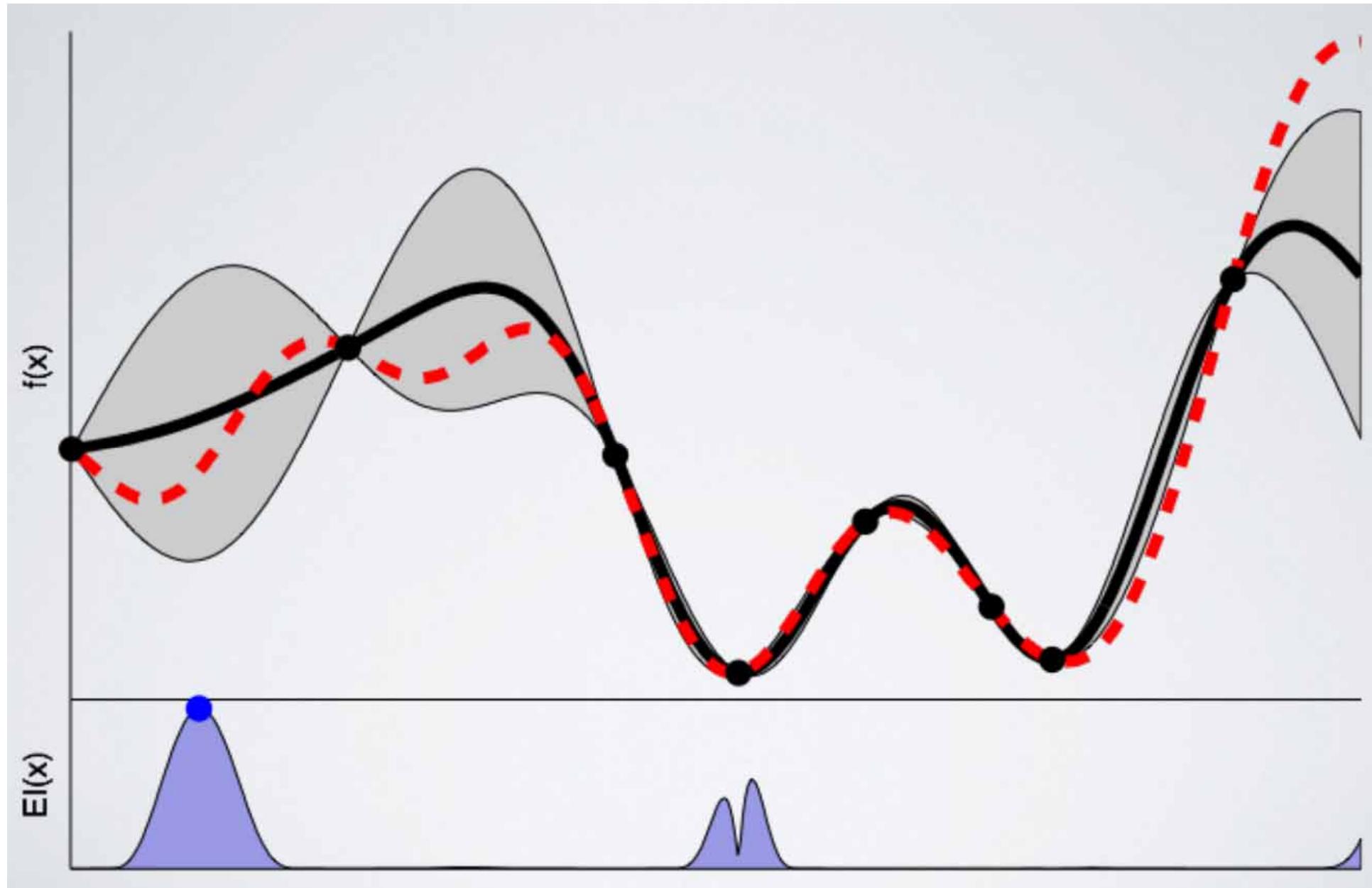












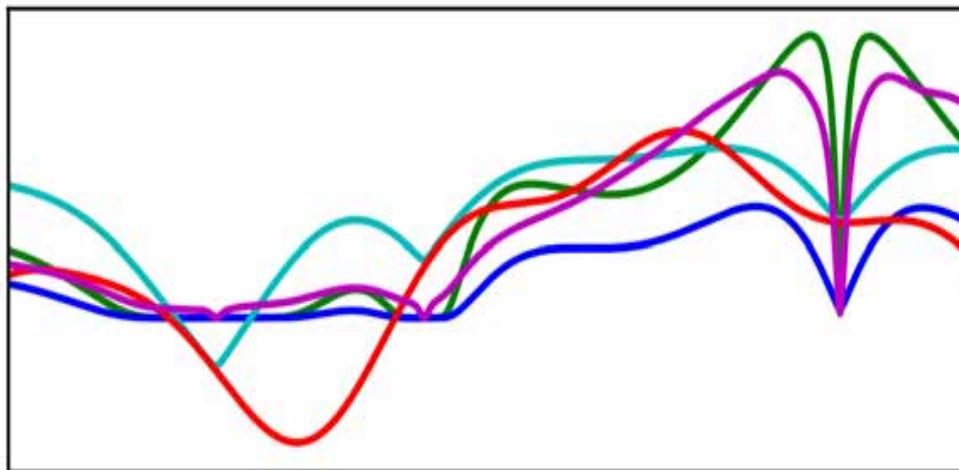
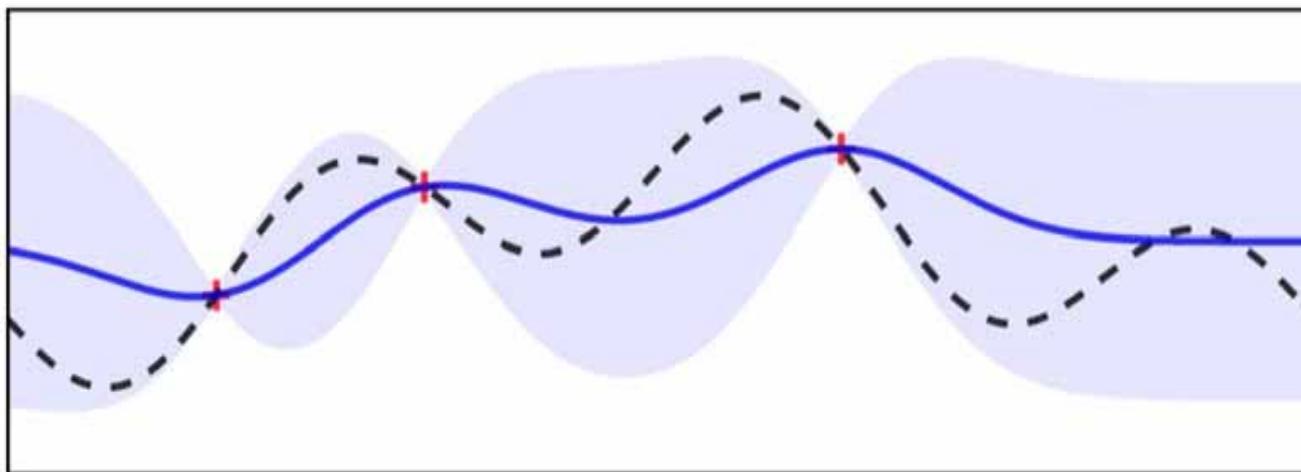
Why is this relevant for health informatics?

- Take patient information, e.g., observations, symptoms, test results, -omics data, etc. etc.
- Reach conclusions, and **predict** into the future, e.g. how likely will the patient be ...
- Prior = belief before making a particular observation
- Posterior = belief after making the observation and is the prior for the next observation – intrinsically incremental

$$p(x_i|y_j) = \frac{p(y_j|x_i)p(x_i)}{\sum p(x_i, y_j)p(x_i)}$$

Grand Goal of automatic Machine Learning

...



Algorithm 1: Bayesian optimization

```
1: for  $n = 1, 2, \dots$ , do
2:   select new  $\mathbf{x}_{n+1}$  by optimizing acquisition function  $\alpha$ 
      
$$\mathbf{x}_{n+1} = \arg \max_{\mathbf{x}} \alpha(\mathbf{x}; \mathcal{D}_n)$$

3:   query objective function to obtain  $y_{n+1}$ 
4:   augment data  $\mathcal{D}_{n+1} = \{\mathcal{D}_n, (\mathbf{x}_{n+1}, y_{n+1})\}$ 
5:   update statistical model
6: end for
```

Shahriari, B., Swersky, K., Wang, Z., Adams, R. P. & De Freitas, N. 2016.
Taking the human out of the loop: A review of Bayesian optimization.
Proceedings of the IEEE, 104, (1), 148-175, doi:10.1109/JPROC.2015.2494218.

- Today most ML-applications are using automatic Machine Learning (aML) approaches
- automatic Machine Learning (aML)
:= algorithms which interact with agents and can optimize their learning behaviour through this interaction

Best practice examples of aML



Wenger 600638 IBEX 17" Laptop Backpack with
Tablet / eReader Pocket (Black / Blue)
von Wenger

EUR 66,99

Andere Angebote

EUR 60,00 neu (22 Angebote)

EUR 57,70 gebraucht (1 Angebot)

Nur Artikel von Wenger anzeigen

★★★★★ 295



Für größere Ansicht Maus über das Bild ziehen

Lenovo Einsteiger Notebooks mit 15,6" HD-Auflösung, 4GB Arbeitsspeicher und Windows 10 Home von Lenovo

★★★★★ 70 Kundenrezensionen

Bestseller Nr. 1 in Notebooks

Unverb. Preisempf.: EUR 349,00

Preis: EUR 273

Sie sparen: EUR 75,01

Alle Preisanfragen

Lieferung Mittwoch, 6. Juli: Bevorzugt
Details.

Auf Lager.

Verkauf und Versand durch Amazon

20 neu ab EUR 273,99 4 gebraucht ab EUR 273,99

Größe: 500GB

1TB

500GB

Stil: Intel Pentium

Intel Core i3

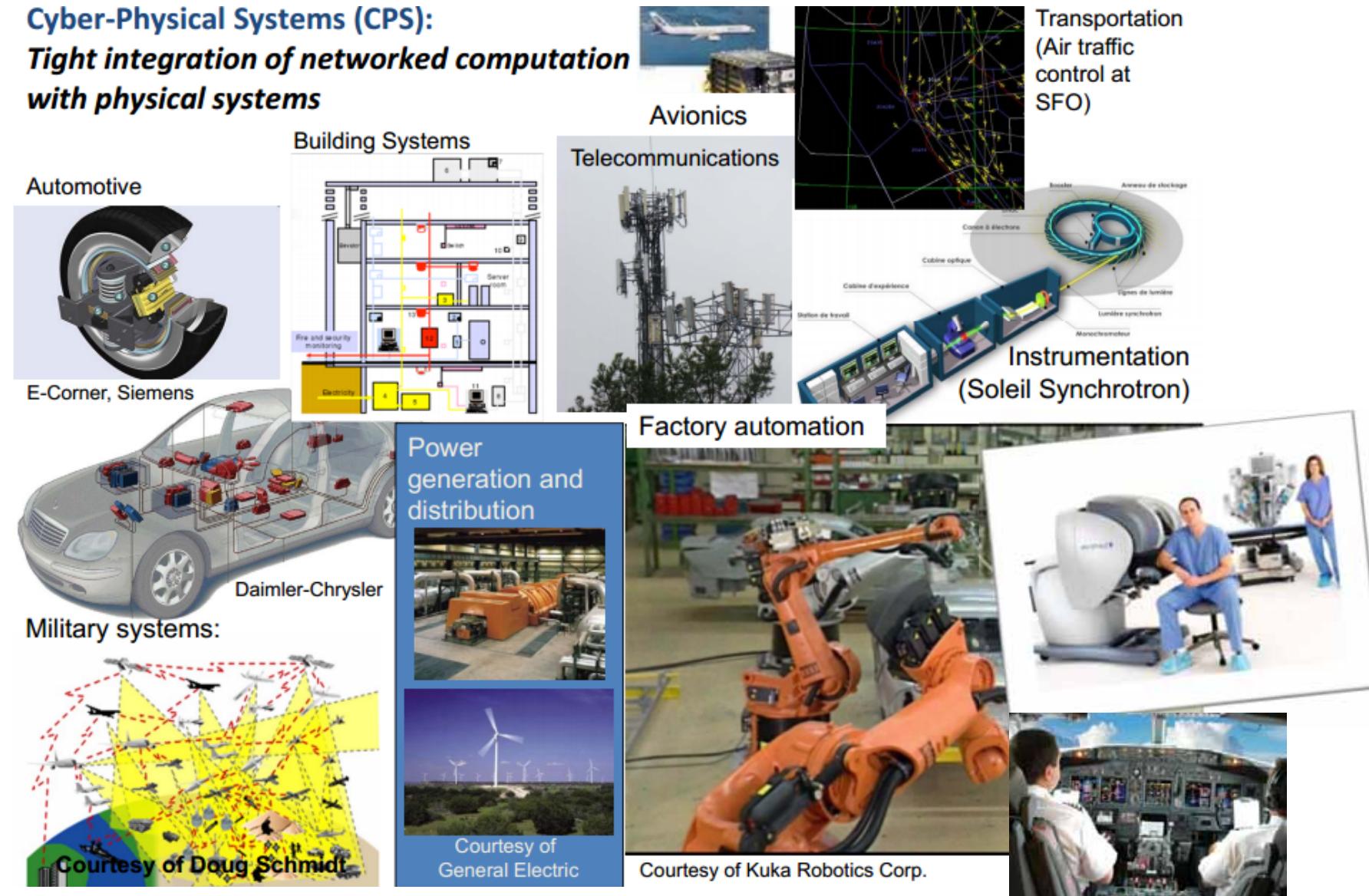
Intel Pentium

- Prozessor: Intel Pentium N35
 - Besonderheiten: HD Glare Display
 - Akku: bis zu 4 Stunden Akkulaufzeit
 - Herstellergarantie: 12 Monate
 - Angaben des jeweiligen Verkäufers
 - Lieferumfang: Lenovo ideapad 300-15IBR
- [Weitere Produktdetails](#)



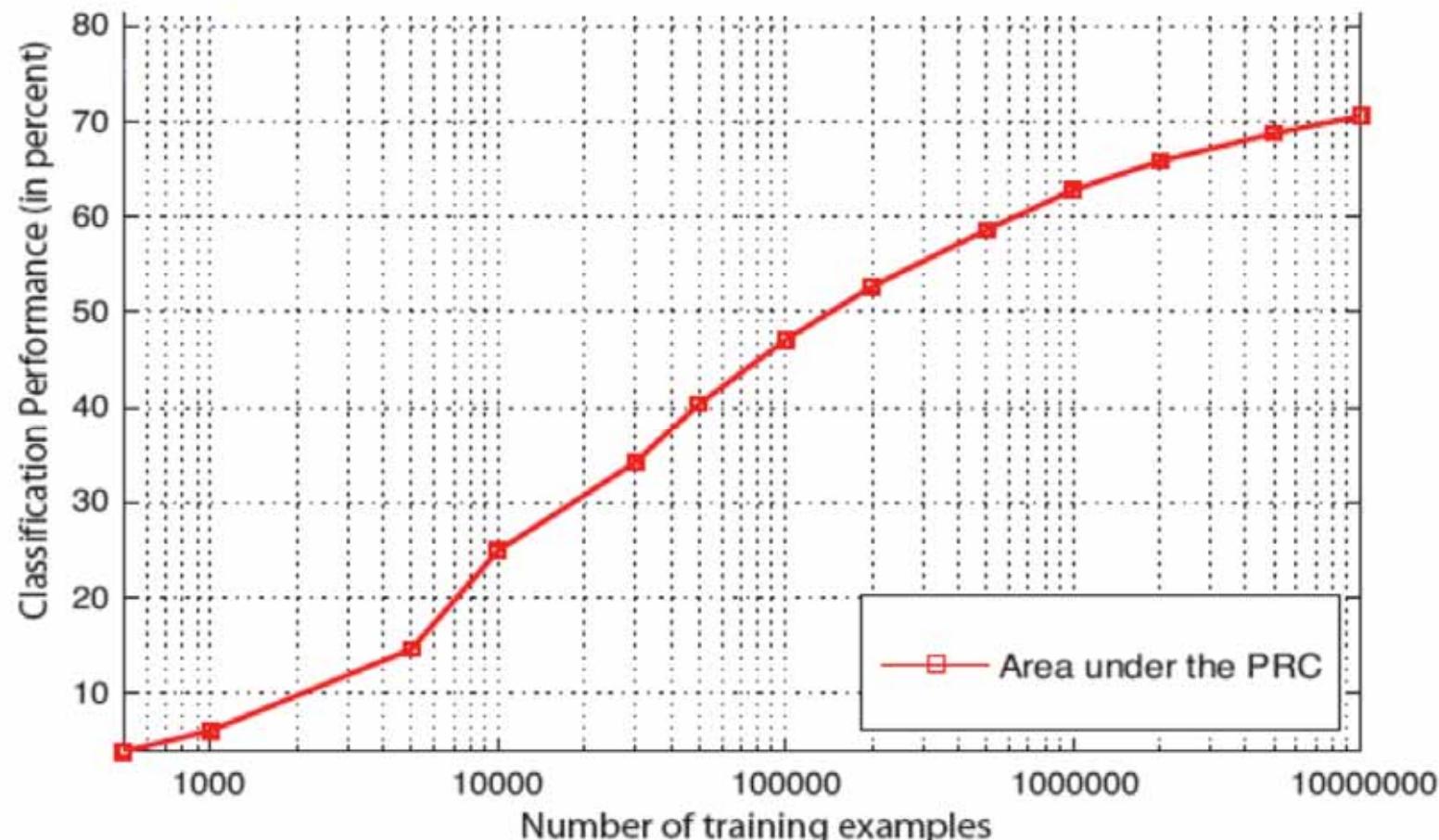
Dietterich, T. G. & Horvitz, E. J. 2015. Rise of concerns about AI: reflections and directions. Communications of the ACM, 58, (10), 38-40.

Cyber-Physical Systems (CPS):
*Tight integration of networked computation
 with physical systems*



Seshia, S. A., Juniwal, G., Sadigh, D., Donze, A., Li, W., Jensen, J. C., Jin, X., Deshmukh, J., Lee, E. & Sastry, S. 2015. Verification by, for, and of Humans: Formal Methods for Cyber-Physical Systems and Beyond. Illinois ECE Colloquium.

Big Data is necessary for aML !



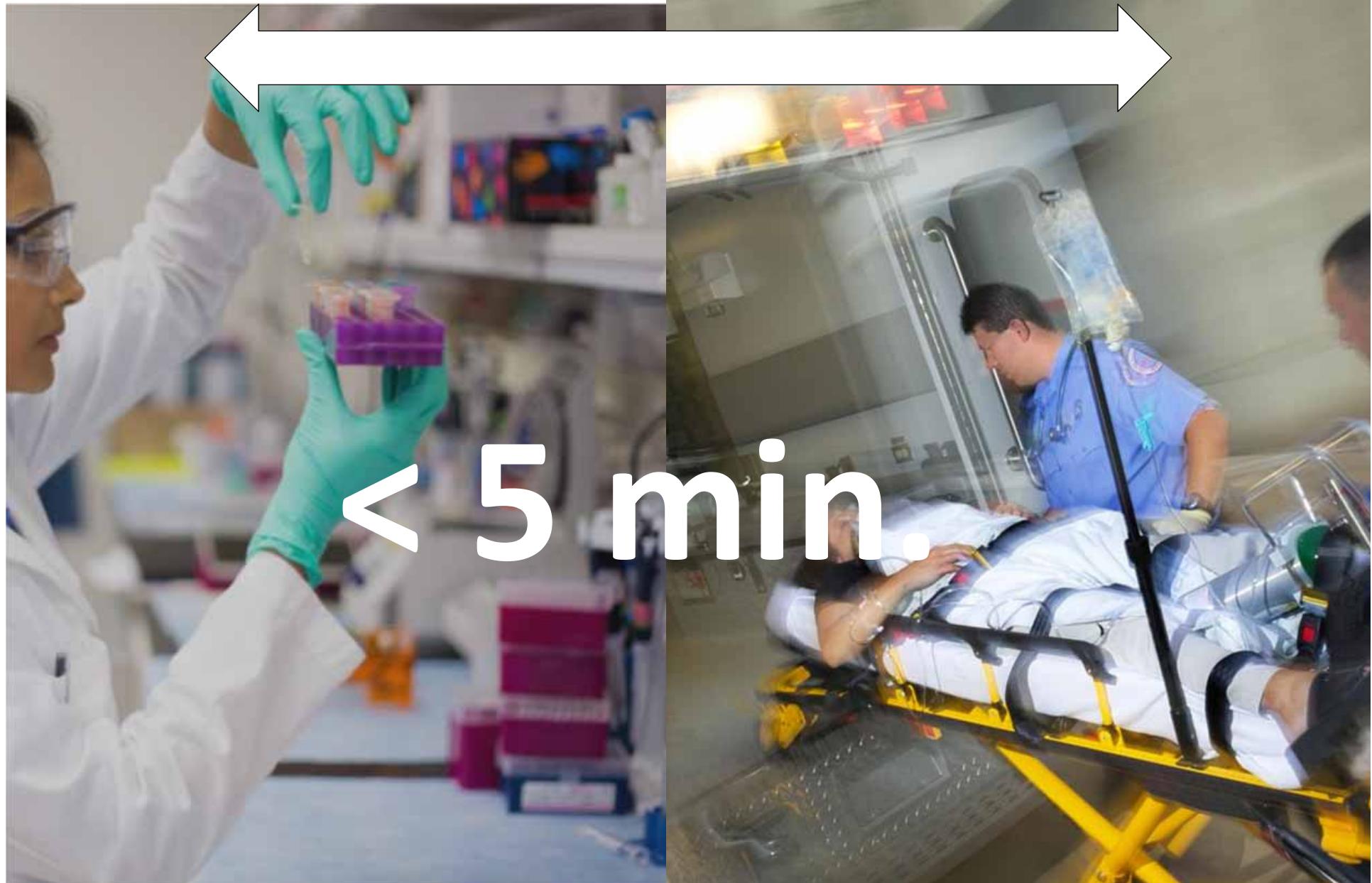
Sonnenburg, S., Rätsch, G., Schäfer, C. & Schölkopf, B. 2006. Large scale multiple kernel learning. Journal of Machine Learning Research, 7, (7), 1531-1565.



Medical Decision Making as a Search Task in \mathcal{H} Problem: Time (t)

Search in an arbitrarily high-dimensional space < 5 min. !

HCI-KDD



- Sometimes we **do not have “big data”**, where aML-algorithms benefit.
- Sometimes we have
 - **Small amount of data sets**
 - Rare Events – **no training samples**
 - **NP-hard problems**, e.g.
 - Subspace Clustering,
 - Protein-Folding,
 - k-Anonymization,
 - Graph Coloring, Category Discovery, etc. etc....

Hans Holbein d.J., 1533,
The Ambassadors,
London: National Gallery

Lopez-Paz, D., Muandet, K., Schölkopf, B. & Tolstikhin, I. 2015. Towards a learning theory of cause-effect inference. Proceedings of the 32nd International Conference on Machine Learning, JMLR, Lille, France.



<https://www.youtube.com/watch?v=9KiVNIUMmCc>

**Sometimes we
(still) need a
human-in-the-loop**

A doctor-in-the-loop



- **interactive Machine Learning (iML)**
:= algorithms which interact with agents*) and can optimize their learning behaviour through this interaction
- ***)where the agents can be human**

Holzinger, A. 2015. Interactive Machine Learning (iML). Informatik Spektrum
DOI: 10.1007/s00287-015-0941-6

A group of experts-in-the-loop



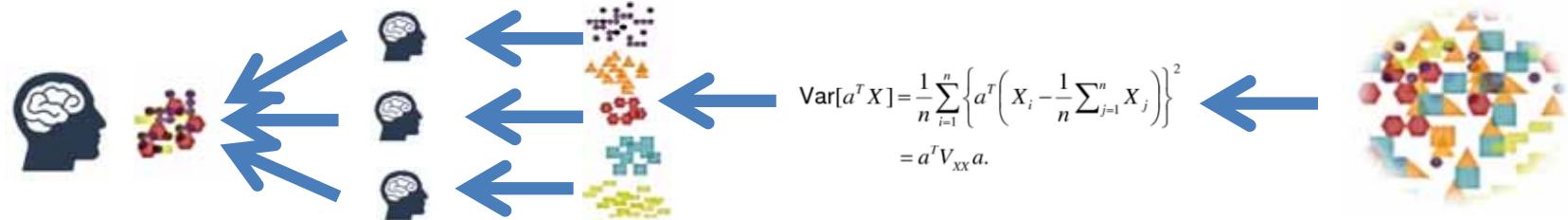
A crowd of people-in-the-loop



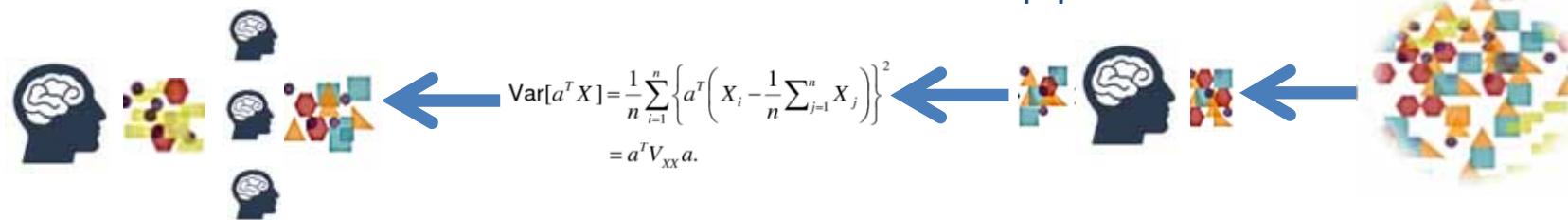




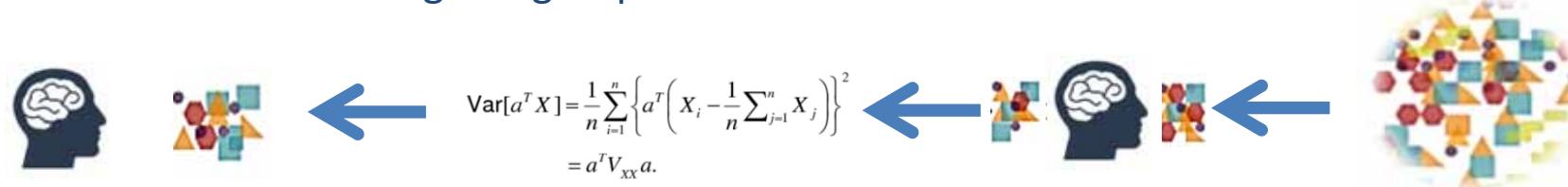
A) Unsupervised ML: Algorithm is applied on the raw data and learns fully automatic – Human can check results at the end of the ML-pipeline



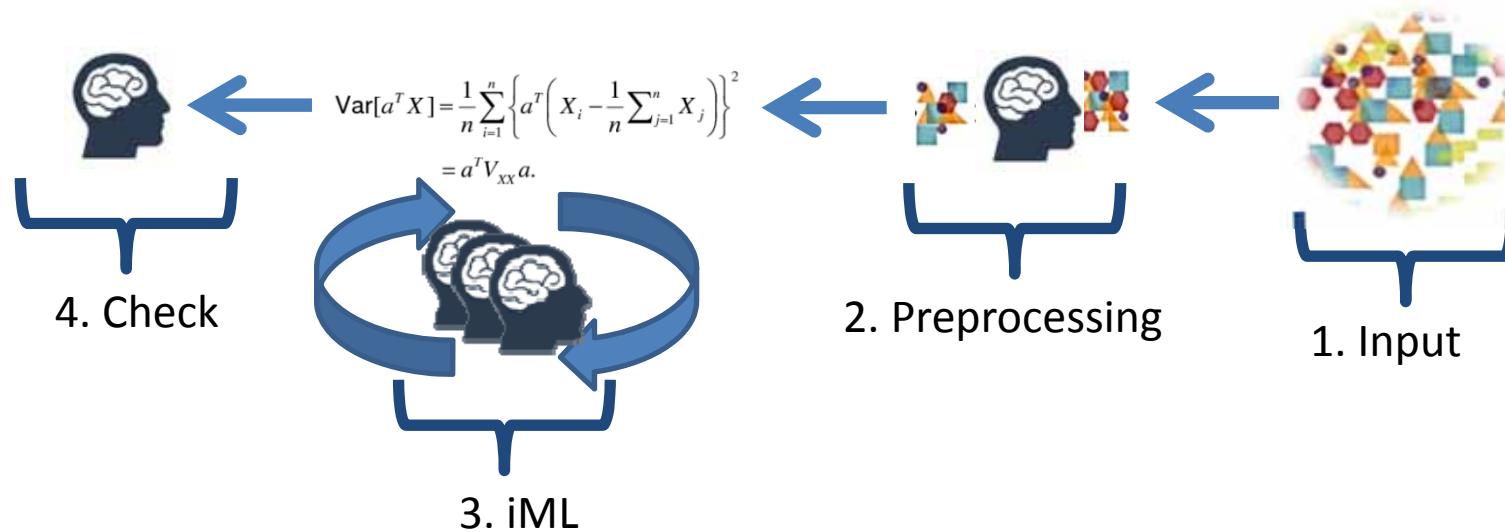
B) Supervised ML: Humans are providing the labels for the training data and/or select features to feed the algorithm to learn – the more samples the better – Human can check results at the end of the ML-pipeline



C) Semi-Supervised Machine Learning: A mixture of A and B – mixing labeled and unlabeled data so that the algorithm can find labels according to a similarity measure to one of the given groups



D) Interactive Machine Learning: Human is seen as an agent involved in the actual learning phase, step-by-step influencing measures such as distance, cost functions ...



Constraints of humans: Robustness, subjectivity, transfer?
Open Questions: Evaluation, replicability, ...

Holzinger, A. 2016. Interactive Machine Learning for Health Informatics: When do we need the human-in-the-loop? *Brain Informatics (BRIN)*, 3, (2), 119-131, doi:10.1007/s40708-016-0042-6.

- Example 1: k-Anonymity
- Example 2: Protein Folding
- Example 3: Subspace Clustering

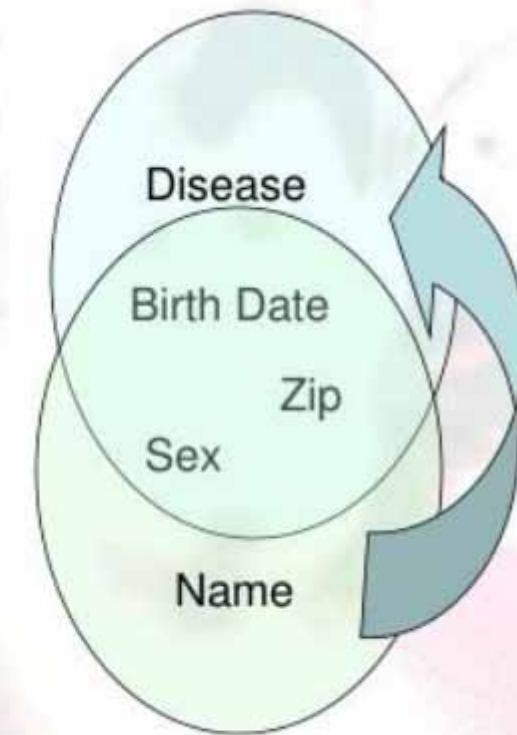
87 % of the population in the USA can be uniquely re-identified by Zip-Code, Gender and date of birth

Hospital Patient Data

Birthdate	Sex	Zipcode	Disease
1/21/76	Male	53715	Flu
4/13/86	Female	53715	Hepatitis
2/28/76	Male	53703	Brochitis
1/21/76	Male	53703	Broken Arm
4/13/86	Female	53706	Sprained Ankle
2/28/76	Female	53706	Hang Nail

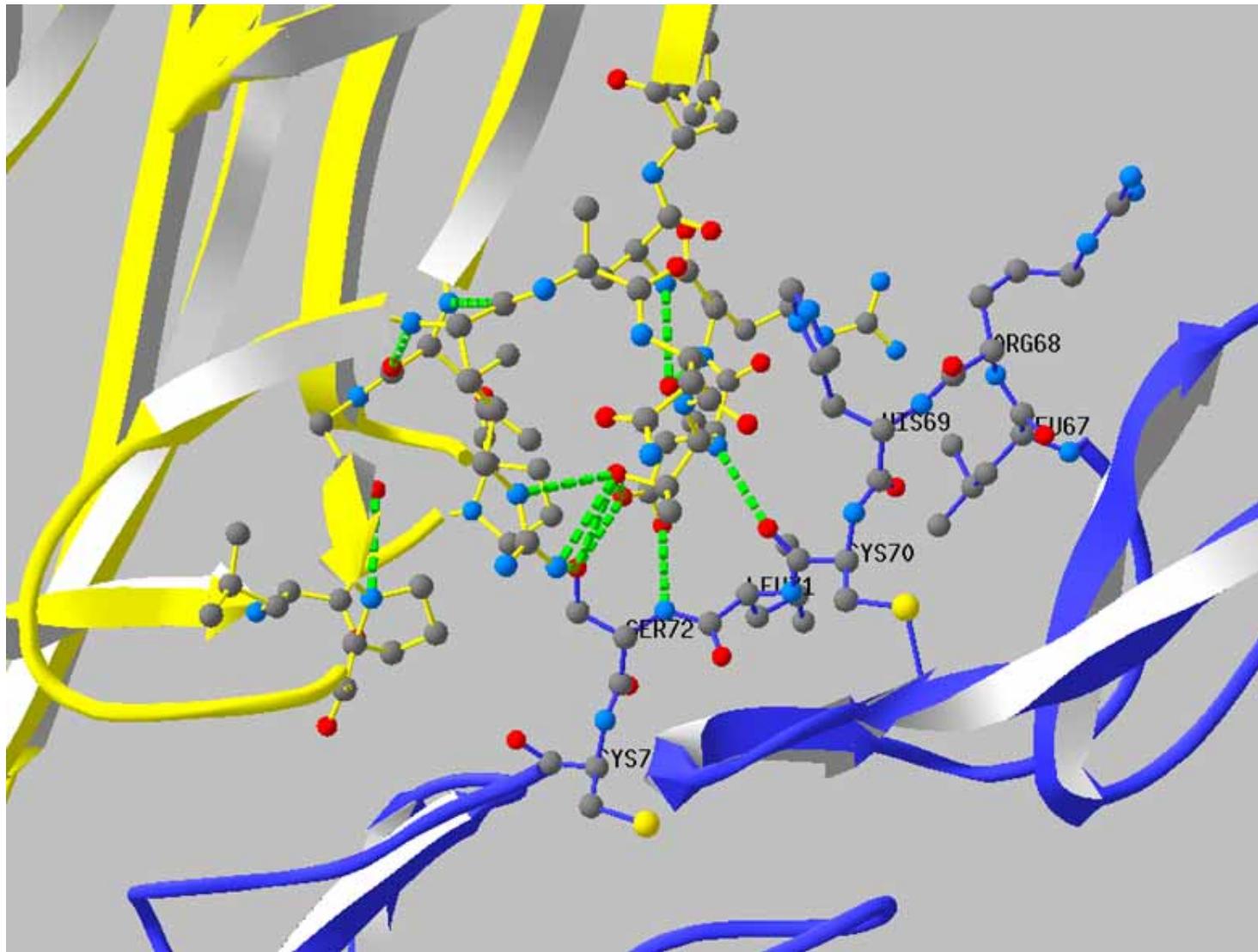
Voter Registration Data

Name	Birthdate	Sex	Zipcode
Andre	1/21/76	Male	53715
Beth	1/10/81	Female	55410
Carol	10/1/44	Female	90210
Dan	2/21/84	Male	02174
Eller	4/19/72	Female	02237



Sweeney, L. 2002. Achieving k-anonymity privacy protection using generalization and suppression. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 10, (05), 571-588.

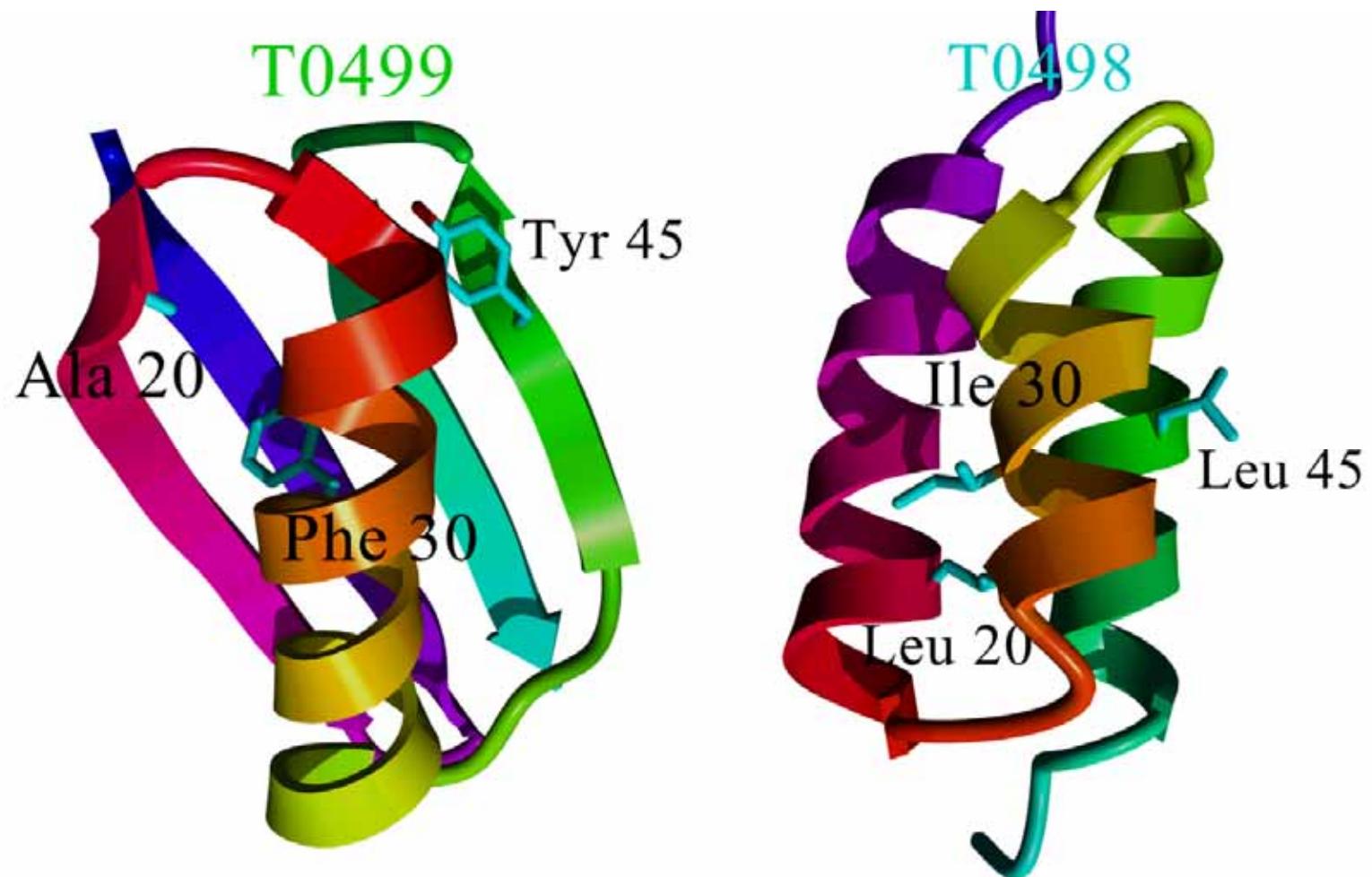
Proteins are the building blocks of life ...



Wiltgen, M., Holzinger, A. & Tilz, G. P. (2007) Interactive Analysis and Visualization of Macromolecular Interfaces Between Proteins. In: *Lecture Notes in Computer Science (LNCS 4799)*. Berlin, Heidelberg, New York, Springer, 199-212.

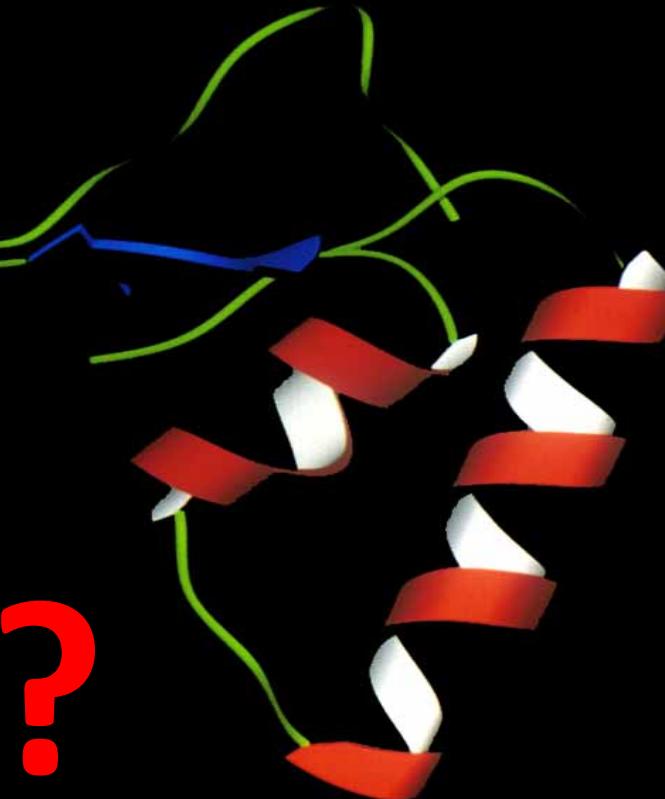
Example 2 Protein Folding: όμολογέω (homologeo)

He, Y., Chen, Y.,
Alexander, P.,
Bryan, P. N. &
Orban, J. (2008)
NMR structures of
two designed
proteins with high
sequence identity
but different fold
and function.
Proceedings of the
National Academy
of Sciences, 105,
38, 14412.



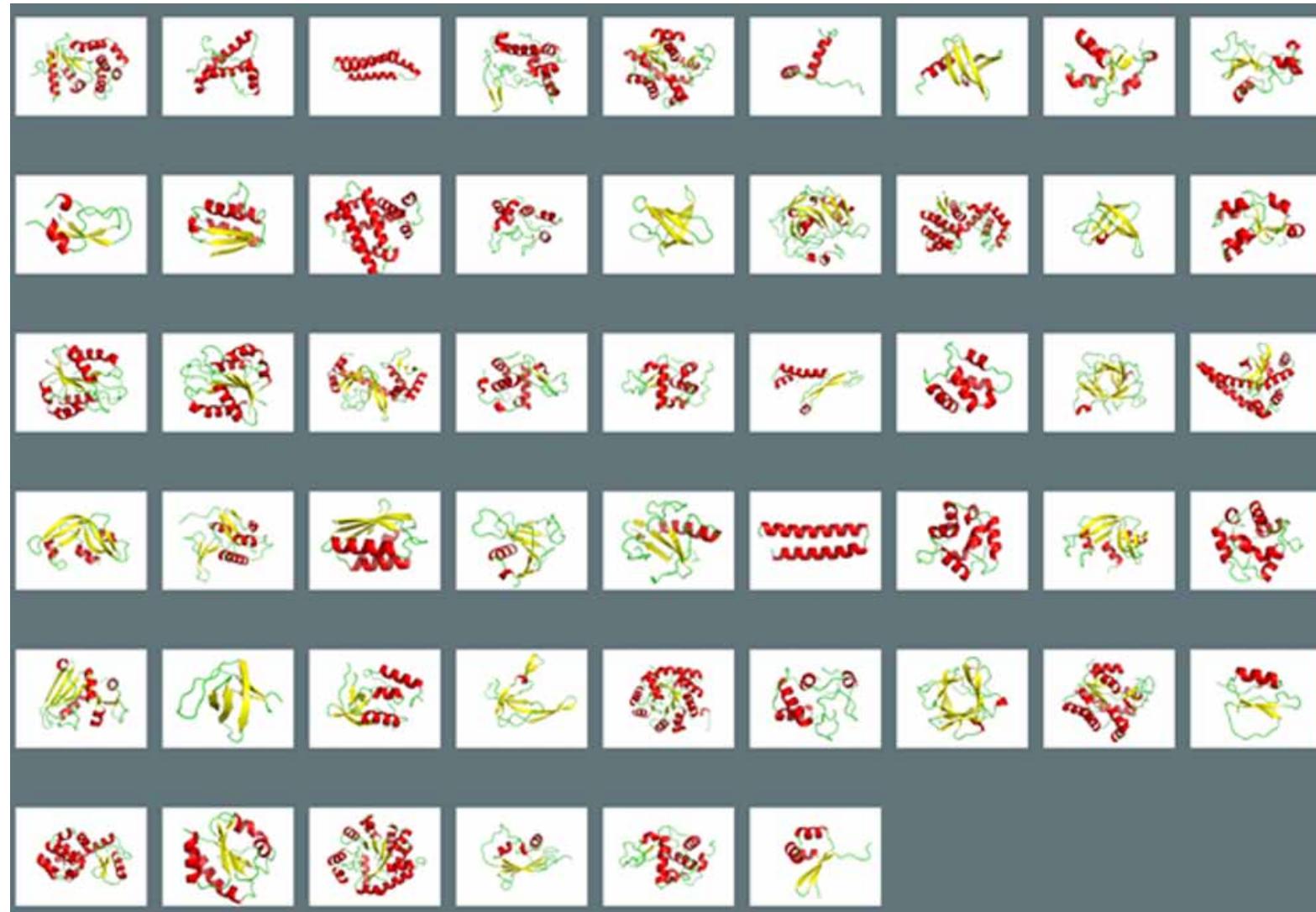
T0499	TTYKL LNL KQAKEEA KEAVDAGTAEKYFKL ANAKTVEGVWTYKDE KTFTVTE
	X X X
T0498	TTYKL LNL KQAKEEA KELVDAGTAEKY KLI ANAKTVEGVWTLKDE KTFTVTE
	X X X

The sequence
of a protein
can NOT (yet)
be used to
predict its 3D
structure ...



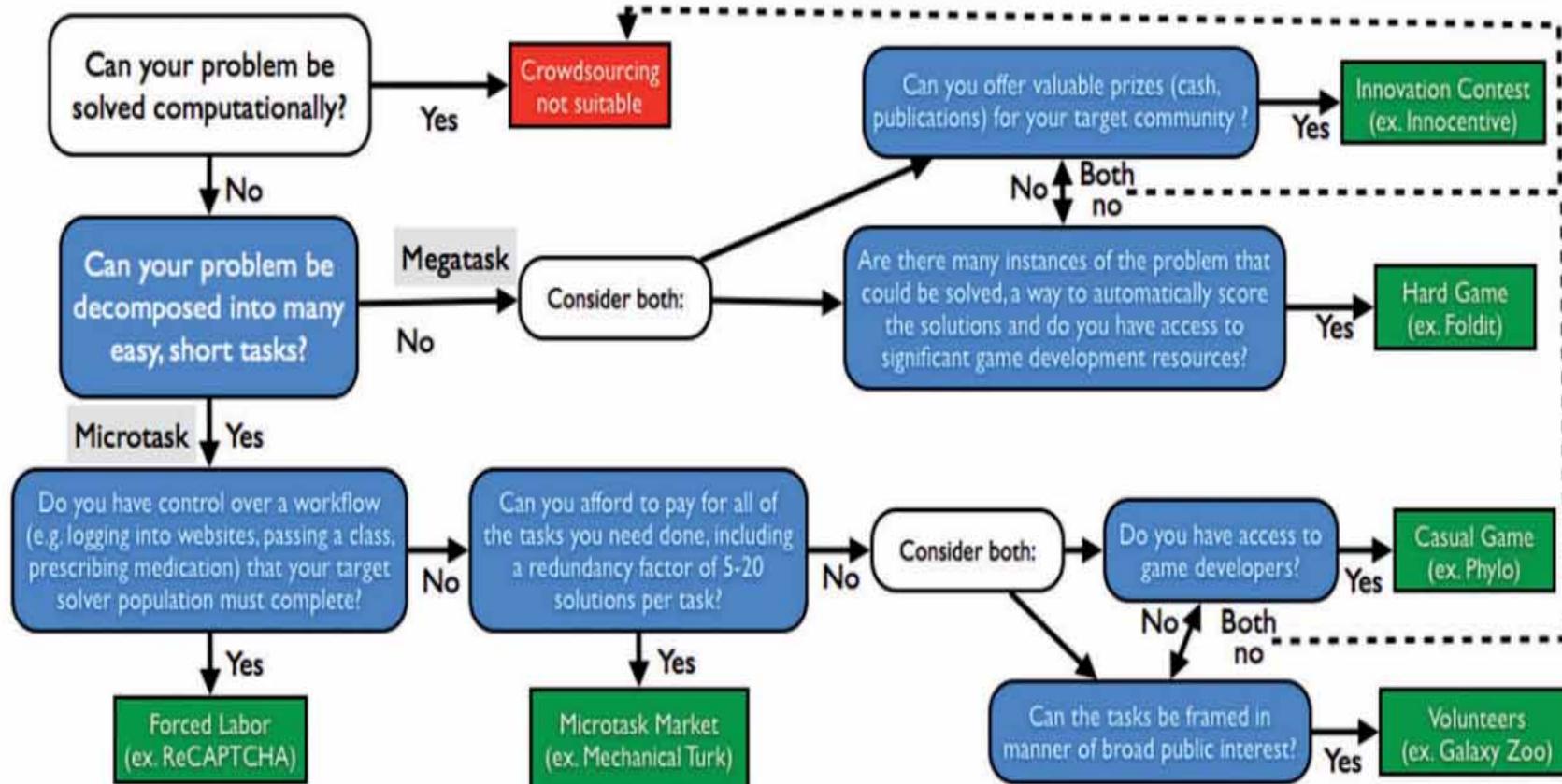
TTCCPSIVARSNFNVCR LPGTPEALCATYTGCIIIPGATCPGDYAN

Anfinsen, C. B. 1973. Principles that Govern the Folding of Protein Chains. *Science*, 181, (4096), 223-230.



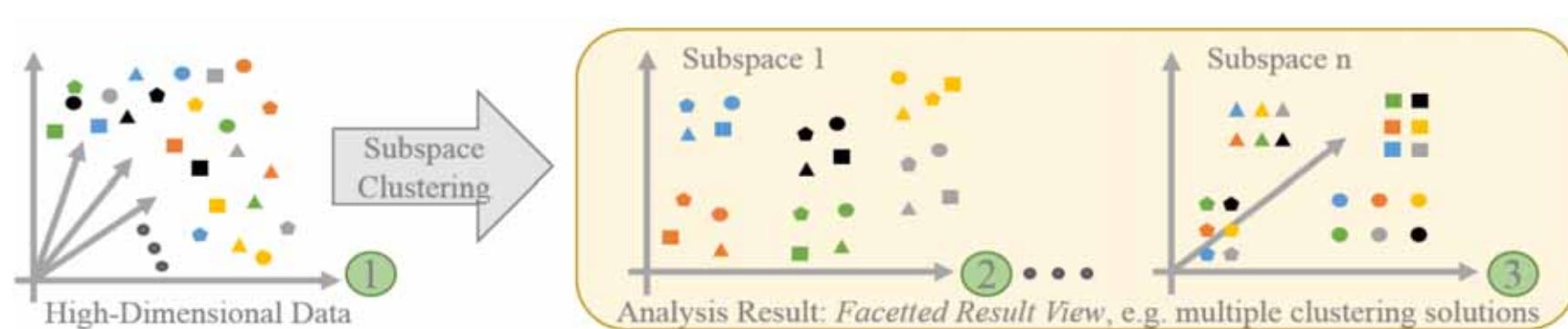
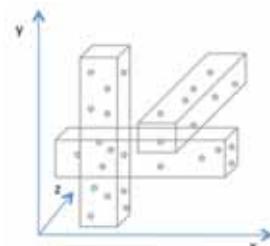
Jia, L., Yarlagadda, R. & Reed, C. C. 2015. Structure Based Thermostability Prediction Models for Protein Single Point Mutations with Machine Learning Tools. Plos One, 10, (9).

Humans can help here: Crowdsourcing

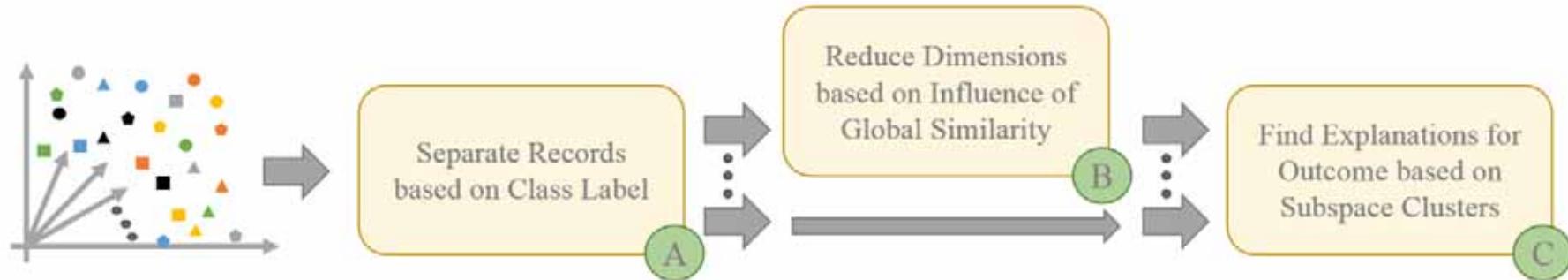


Good, B. M. & Su, A. I. 2013. Crowdsourcing for bioinformatics. *Bioinformatics*, 29, (16), 1925-1933.

- Patterns may be found in subspaces (dimension combinations)
- Clustering and subset selection: Non-convex & NP-hard
- Real data are often noisy and corrupted
- Little prior knowledge about low-dim structures
- Data points in different subgroups can be very close



Hund, M., Sturm, W., Schreck, T., Ullrich, T., Keim, D., Majnaric, L. & Holzinger, A. 2015. Analysis of Patient Groups and Immunization Results Based on Subspace Clustering. In: Lecture Notes in Artificial Intelligence LNAI 9250, pp. 358-368.



■ Positive subspace clusters

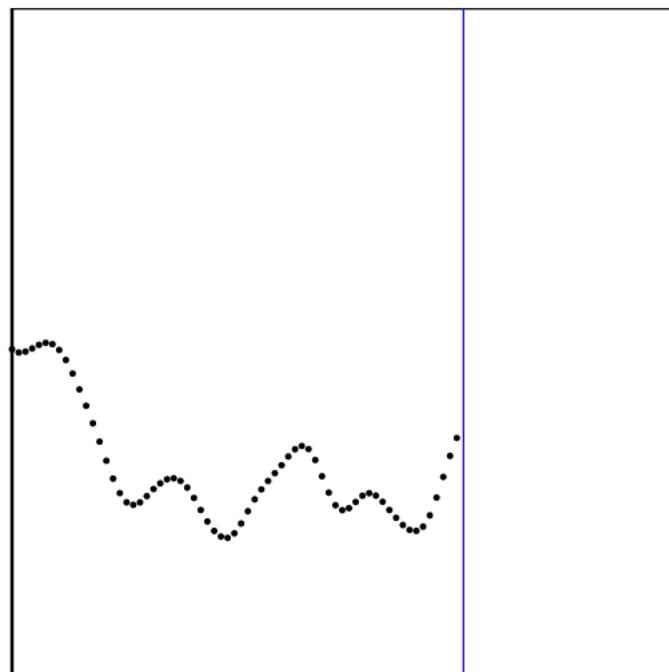
- One homogeneous cluster (healthy patients)
 - hyper, CVD, neoplasm, psy. disorder, drug allergy
 - No medications: statins, anticoagulants, analgesics and clear (preserved renal function)
- Negative subspace clusters
 - Cluster with obvious reasons for neg. outcome
 - Impairment of certain pathophysiologic mechanism increased MCV, decreased VITB12, FOLNA, CORTIS) despite no: DM, drug allergy, Fglu, E/HB (anaemia)

Judgment 1 out of 33

This is the first function from the system. Please try to predict the new points as well as you can based on the points you can see.

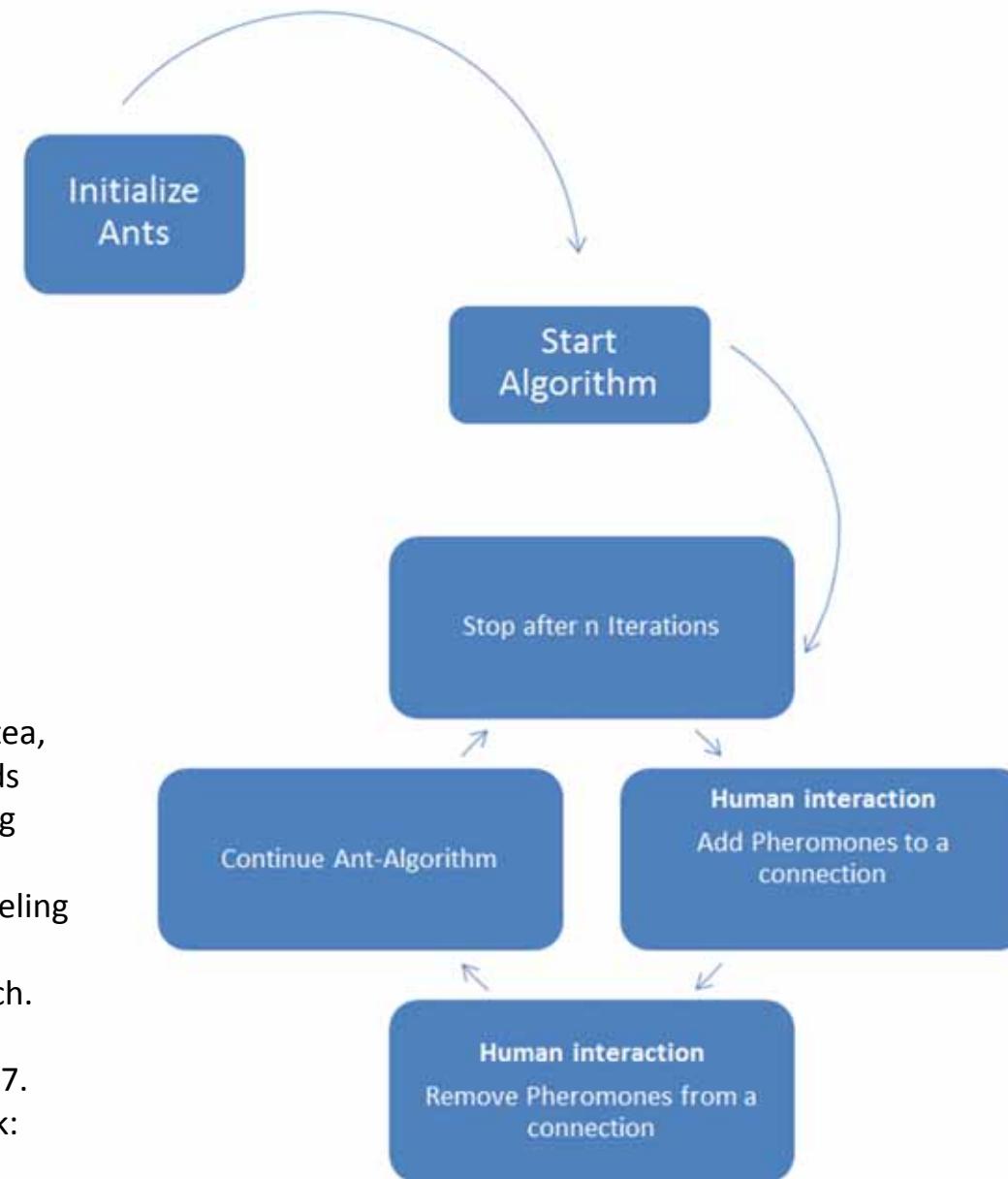
Please click along the blue line to say what you think the height of the point is for that location.

Once you have selected a position along the line, **hit the 's' key to submit the point**.

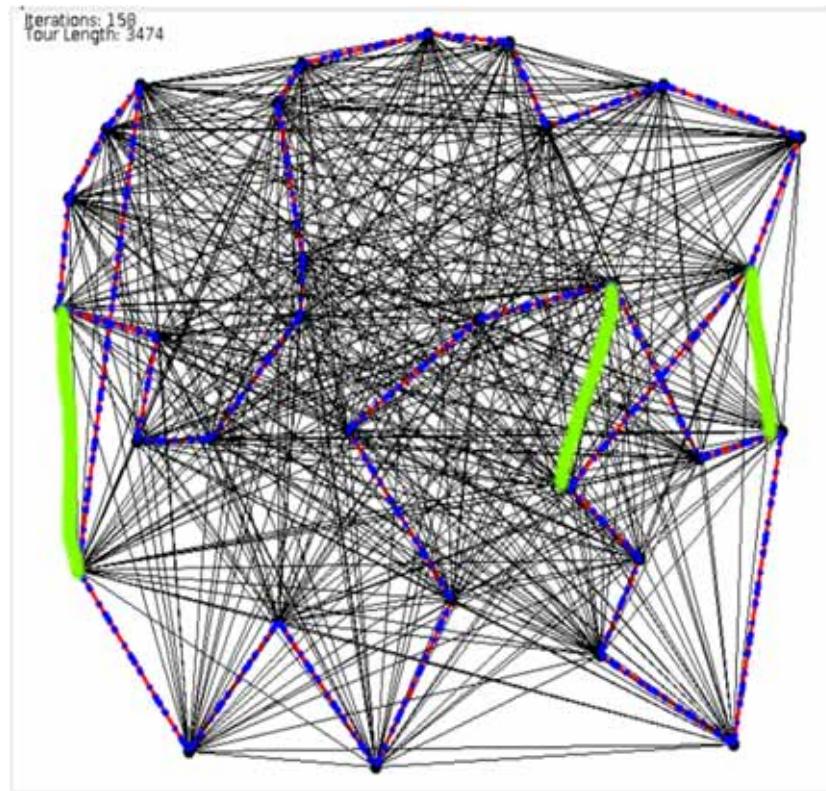


Wilson, A. G., Dann, C., Lucas, C. & Xing, E. P. The Human Kernel. Advances in Neural Information Processing Systems, 2015. 2836-2844.

$$\hat{p}_{bc}^a = \frac{\mu + \delta_{ac}}{2\mu + \delta_{ab} + \delta_{ac}} \text{ and } K_{ii} = 1,$$



Holzinger, A., Plass, M.,
Holzinger, K., Crisan, G., Pintea,
C. & Palade, V. 2016. Towards
interactive Machine Learning
(iML): Applying Ant Colony
Algorithms to solve the Traveling
Salesman Problem with the
Human-in-the-Loop approach.
Springer Lecture Notes in
Computer Science LNCS 9817.
Heidelberg, Berlin, New York:
Springer, pp. in print.


Algorithm 2: Ant Colony Algorithm iML

```

Input : ProblemSize,  $Population_{size}$ ,  $m$ ,  $\rho$ ,  $\beta$ ,  $\sigma$ ,  $q_0$ 
Output:  $P_{best}$ 
 $P_{best} \leftarrow CreateHeuristicSolution(ProblemSize);$ 
 $P_{best\_cost} \leftarrow Cost(S_h);$ 
 $Pheromone_{init} \leftarrow \frac{1.0}{ProblemSize \times P_{best\_const}};$ 
 $Pheromone \leftarrow InitializePheromone(Pheromone_{init});$ 
while  $\neg StopCondition()$  do
    for  $i = 1$  to  $m$  do
         $S_i \leftarrow ConstructSolution(Pheromone, ProblemSize, \beta, q_0);$ 
         $S_{i\_cost} \leftarrow Cost(S_i);$ 
        if  $S_{i\_cost} \leq P_{best\_cost}$  then
             $P_{best\_cost} \leftarrow S_{i\_cost};$ 
             $P_{best} \leftarrow S_i;$ 
        end
         $LocalUpdateAndDecayPheromone(Pheromone, S_i, S_{i\_cost}, \sigma);$ 
    end
     $GlobalUpdateAndDecayPheromone(Pheromone, P_{best}, P_{best\_cost}, \rho);$ 
    while  $isUserInteraction()$  do
         $| GlobalAddAndRemovePheromone(Pheromone, P_{best}, P_{best\_cost}, \rho);$ 
    end
end
return  $P_{best};$ 

```

<http://hci-kdd.org/project/iml/>

Holzinger, A., Plass, M., Holzinger, K., Crisan, G., Pintea, C. & Palade, V. 2016. Towards interactive Machine Learning (iML): Applying Ant Colony Algorithms to solve the Traveling Salesman Problem with the Human-in-the-Loop approach. Springer Lecture Notes in Computer Science LNCS 9817. Heidelberg, Berlin, New York: Springer, pp. in print.

- ① Heterogeneous data sources
 - need for data integration and data fusion
- ② Complexity – reduction of search space
 - combining the best of Human & Computer
- ③ What is interesting? – and relevant!
 - need of effective mapping $\mathbb{R}^N \rightarrow \mathbb{R}^2$
- ④ Clinical time limits “5 Minutes”
 - need of efficient solutions

Holzinger, A. & Jurisica, I. 2014. Knowledge Discovery and Data Mining in Biomedical Informatics: The future is in Integrative, Interactive Machine Learning Solutions In: LNCS 8401. Heidelberg, Berlin: Springer, pp. 1-18.

Multi-Task Learning (MTL)

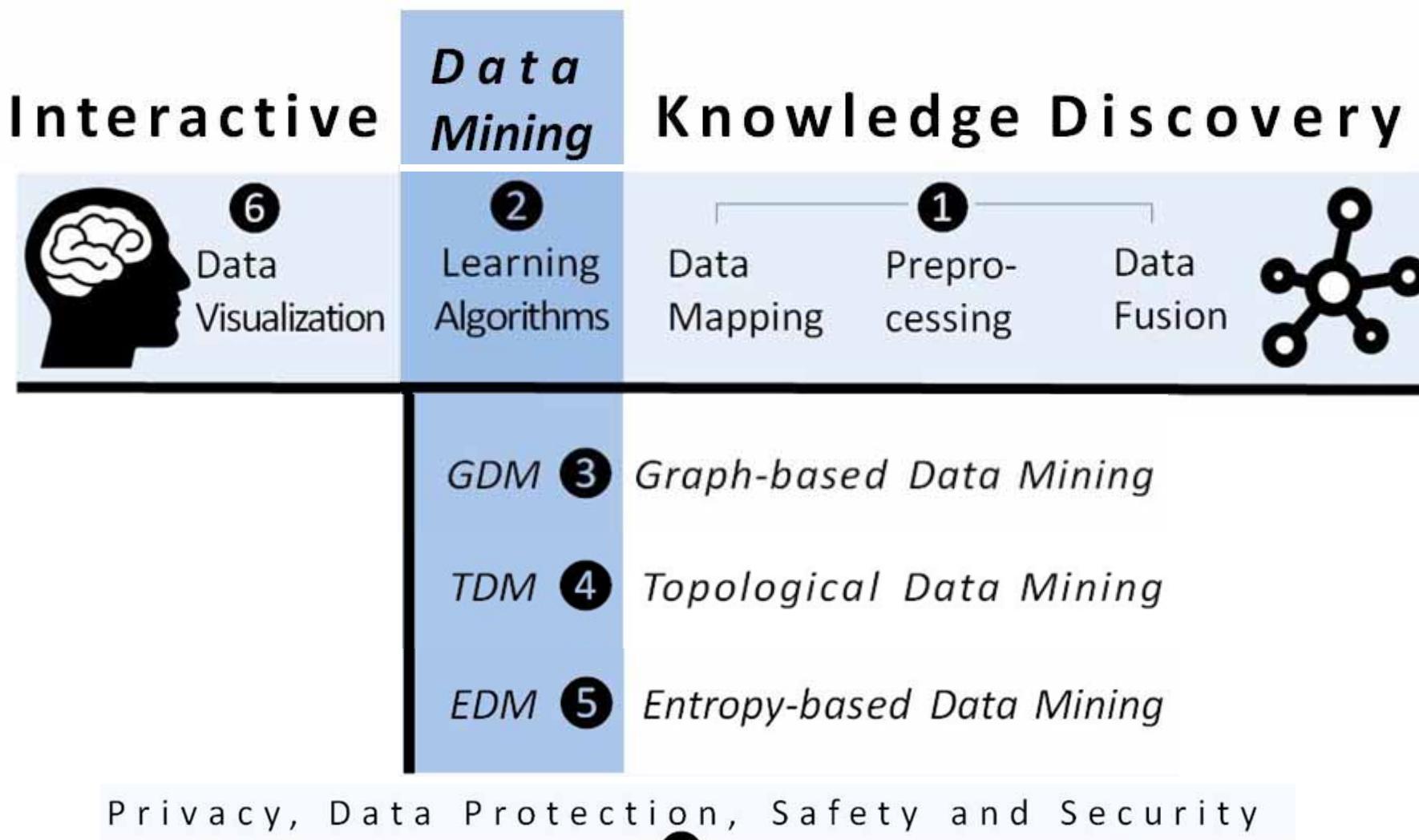
for improving prediction performance, help to reduce
catastrophic forgetting

Transfer learning (TL)

is not easy: learning to perform a task by exploiting
knowledge acquired when solving previous tasks:
a solution to this problem would have major impact to
AI research generally and ML specifically.

Multi-Agent-Hybrid Systems (MAHS)

To include swarm-intelligence and crowdsourcing
and making use of discrete models – avoiding to seek
perfect solutions – better have a good solution < 5 min.



Holzinger, A. 2014. Trends in Interactive Knowledge Discovery for Personalized Medicine: Cognitive Science meets Machine Learning. *IEEE Intelligent Informatics Bulletin*, 15, (1), 6-14.



**concerted effort
international
without boundaries ...**





Thank you!