Seminar Explainable AI

Module 7

Selected Methods Part 3
Feature Vis, Deep Vis, RNN, Fitted Additive and Interactive ML with the human-in-the-loop

Andreas Holzinger

Human-Centered AI Lab (Holzinger Group)

Institute for Medical Informatics/Statistics, Medical University Graz, Austria

and

Explainable AI-Lab, Alberta Machine Intelligence Institute, Edmonton, Canada

This is the version for printing and reading. The lecture version is didactically different.

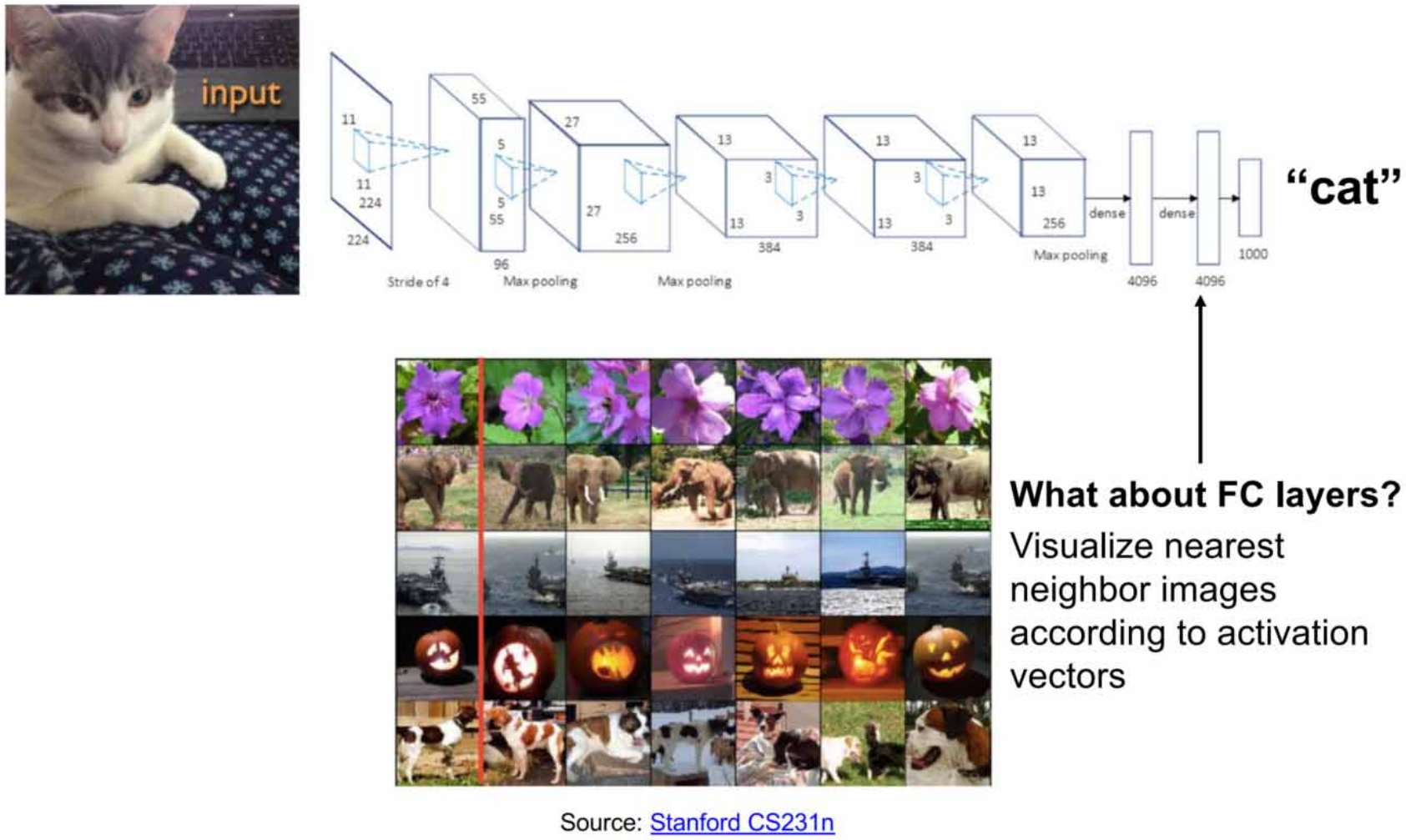
AGENDA

- 00 Reflection
- 01 Feature Visualization
- 02 Deep Visualization
- 03 Recursive Neural Networks cell state analysis
- 04 Fitted Additive
- 05 Interactive Machine Learning with the human-in-the-loop

01 Understanding the Model: Feature Visualization

Goal: Understand what is happening inside the black-box?

- For a better understanding you can review briefly lecture 13 of Stanford CS231n: cs231n.stanford.edu/slides/2018/cs231n_2018_lecture13.pdf



Source: Stanford CS231n

https://cs.stanford.edu/people/karpathy/cnnembed/

Maximizing the activation of a unit as an optimization problem

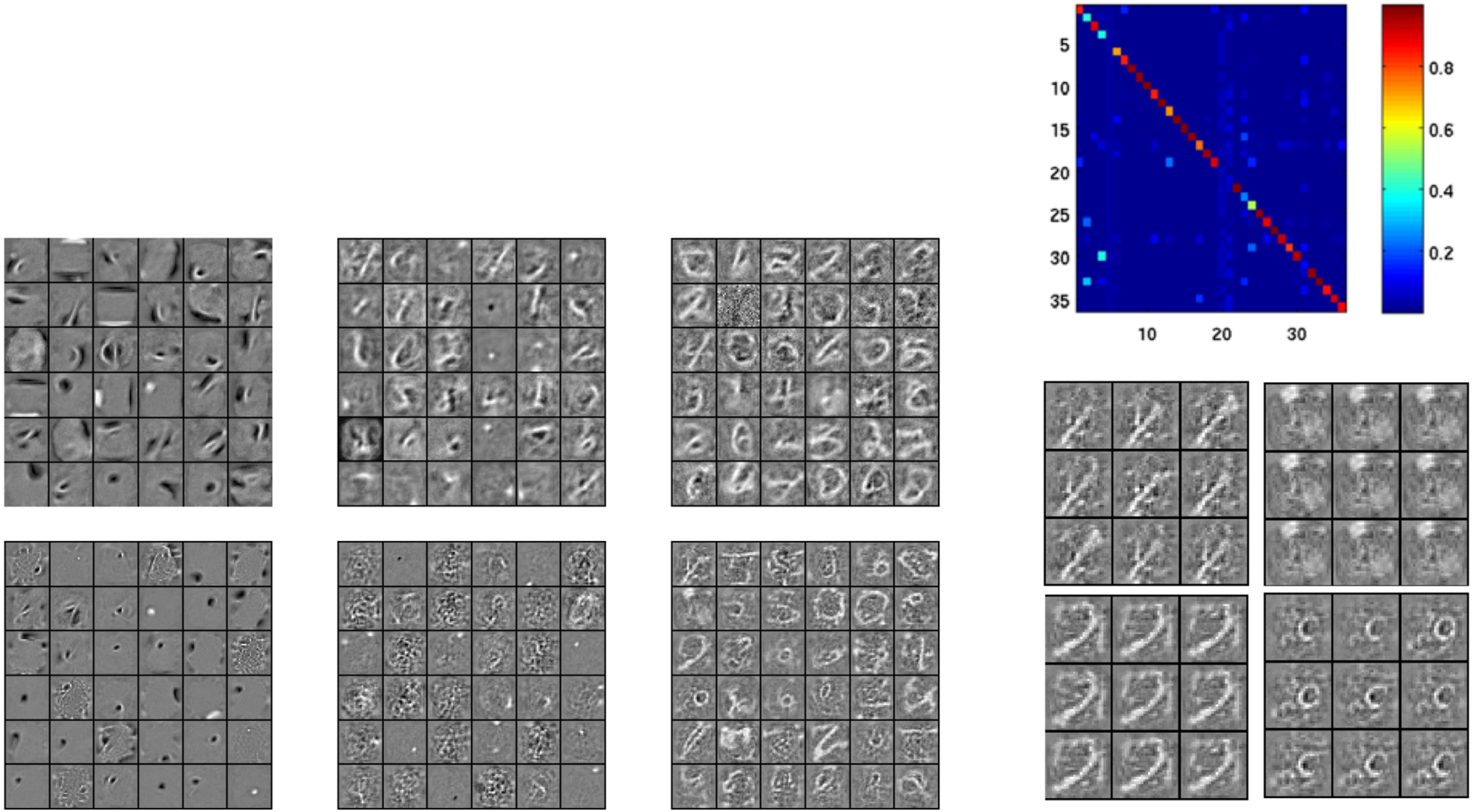
- A pattern to which the unit is responding maximally could be a good first-order representation of what a unit is doing. One simple way of doing this is to find, for a given unit, the input sample(s) (from either the training or the test set) that give rise to the highest activation of the unit.
- Ideally, we would like to find out what these samples have in common. Furthermore, it may be that only some subsets of the input vector **contribute** to the high activation, and it is not easy to determine which by inspection.
- Let θ denote our neural network parameters (weights and biases) and let $h_{ij}(\theta;\ x)$ be the activation of a given unit I from a given layer j in the network; h_{ij} is a function of both θ and the input sample x. Assuming a fixed θ (for instance, the parameters after training the network), we can view this idea as looking for

$$\mathbf{x}^* = \arg \max_{\mathbf{x}\ s.t.\ ||\mathbf{x}||=\rho} h_{ij}(\theta, \mathbf{x})$$

This is, in general, a non-convex optimization problem. But it is a problem for which we can at least try to find a local minimum. This can be done most easily by performing simple gradient ascent in the input space, i.e. computing the gradient of $h_{ij}(\theta;\ x)$ and moving x in the direction of this gradient

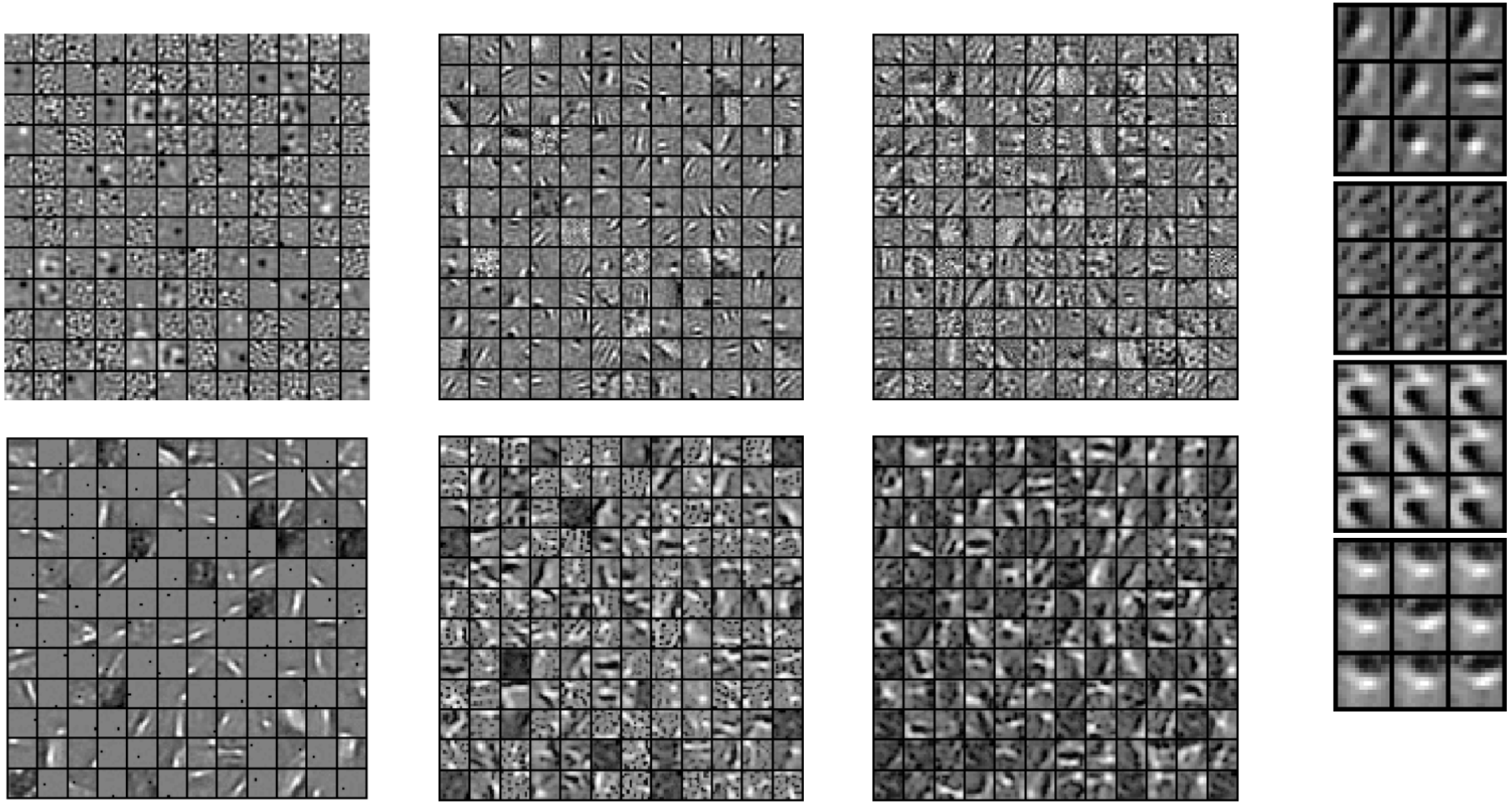
Dumitru Erhan, Yoshua Bengio, Aaron Courville & Pascal Vincent 2009. Visualizing higher-layer features of a deep network. University of Montreal Technical Report Nr. 1341.

Maximizing the activation of a unit as an optimization problem



Dumitru Erhan, Yoshua Bengio, Aaron Courville & Pascal Vincent 2009. Visualizing higher-layer features of a deep network. University of Montreal Technical Report Nr. 1341.

Maximizing the activation of a unit as an optimization problem



Dumitru Erhan, Yoshua Bengio, Aaron Courville & Pascal Vincent 2009. Visualizing higher-layer features of a deep network. University of Montreal Technical Report Nr. 1341.

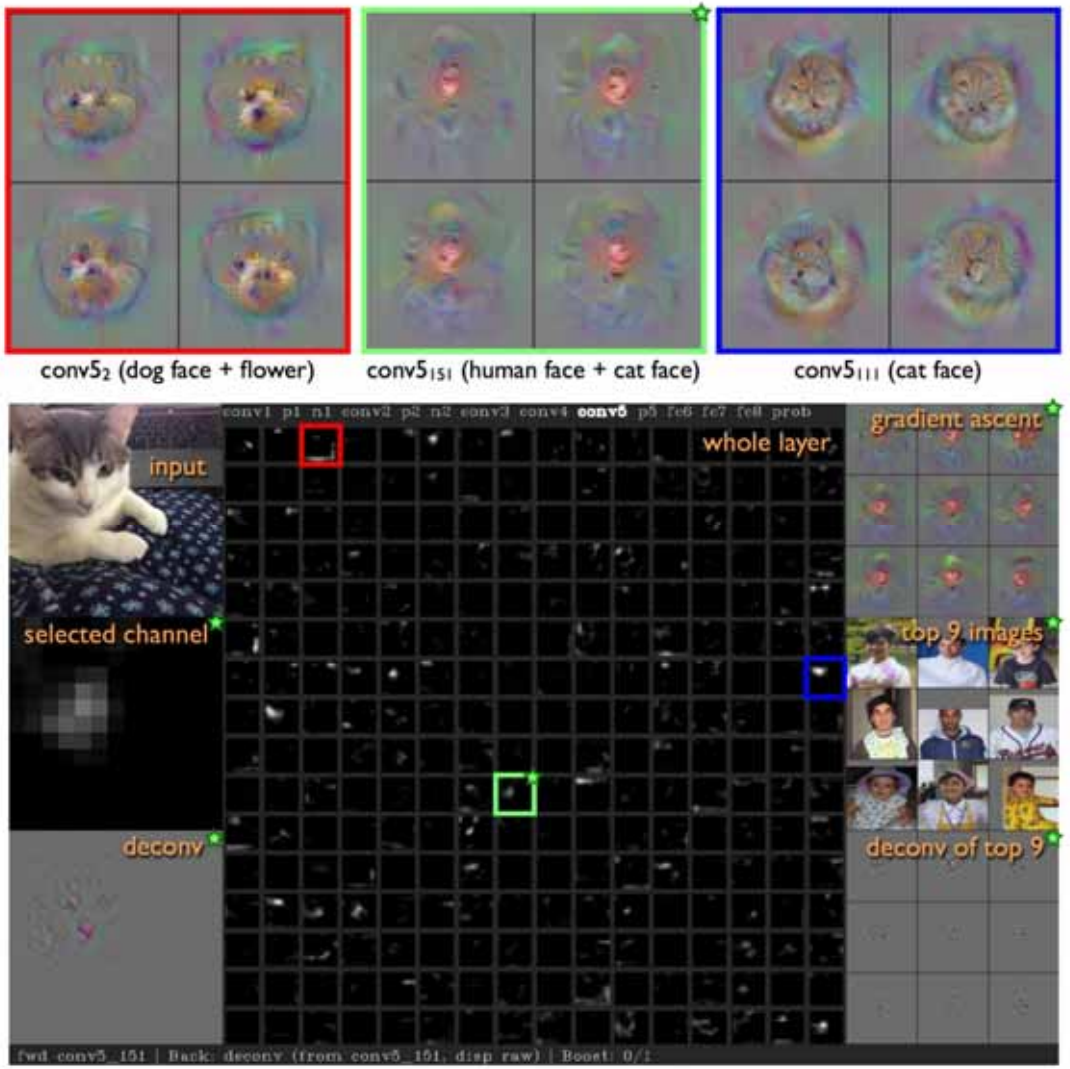
02 Understanding the Model: Deep Visualization

- Alternatively to regularization: at each step of our gradient ascent, apply an operator r which regularizes the image:

$$x \leftarrow r\left(x + \eta \frac{\partial f}{\partial x}\right)$$

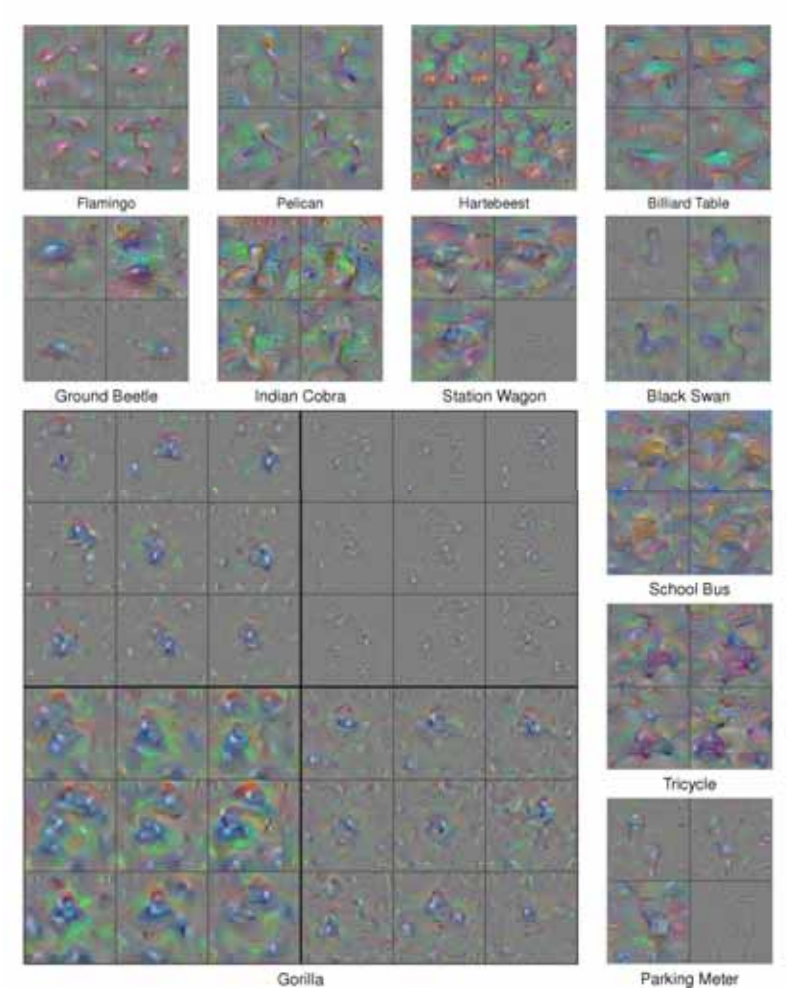
Visualizing conv5 activations

Jason Yosinski, Jeff Clune, Anh Nguyen, Thomas Fuchs & Hod Lipson 2015. Understanding neural networks through deep visualization. arXiv:1506.06579.

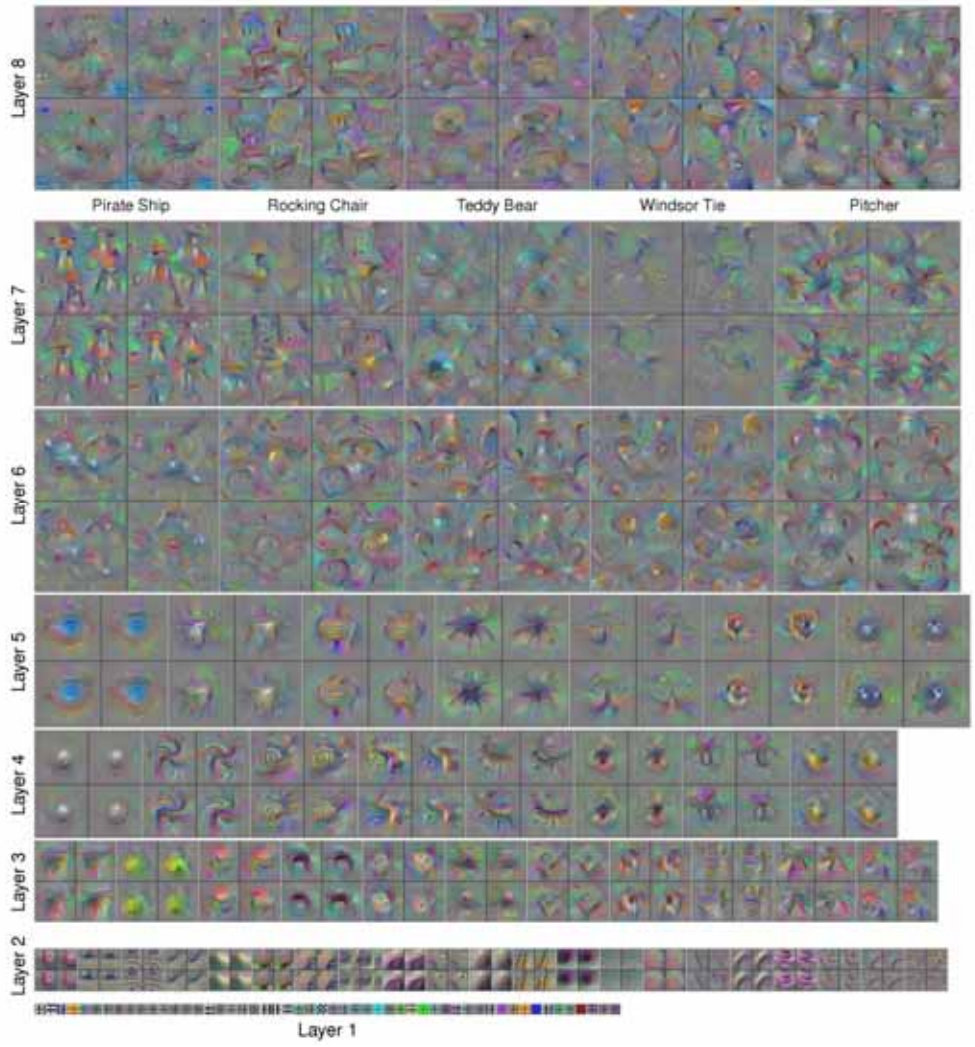


http://yosinski.com/deepvis

Visualization via Regularized Optimization



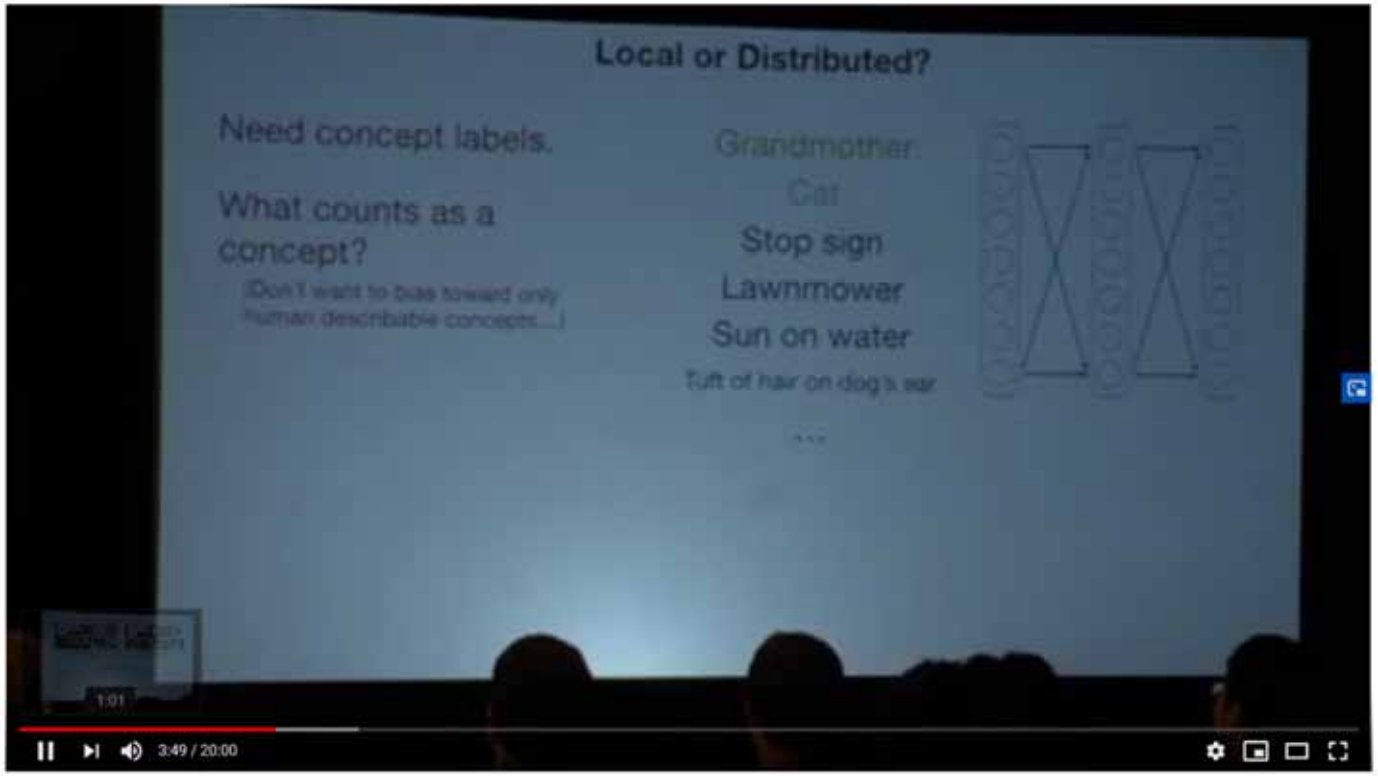
θ_{decay}	$\theta_{\text{b\_width}}$	$\theta_{\text{b\_every}}$	$\theta_{\text{n\_pct}}$	$\theta_{\text{c\_pct}}$
0	0.5	4	50	0
0.3	0	0	20	0
0.0001	1.0	4	0	0
0	0.5	4	0	90



$$\mathbf{x}^* = \arg\max_{\mathbf{x}} (a_i(\mathbf{x}) - R_\theta(\mathbf{x}))$$

Jason Yosinski, Jeff Clune, Anh Nguyen, Thomas Fuchs & Hod Lipson 2015. Understanding neural networks through deep visualization. *arXiv:1506.06579*.

Towards concept learning: local or distributed?



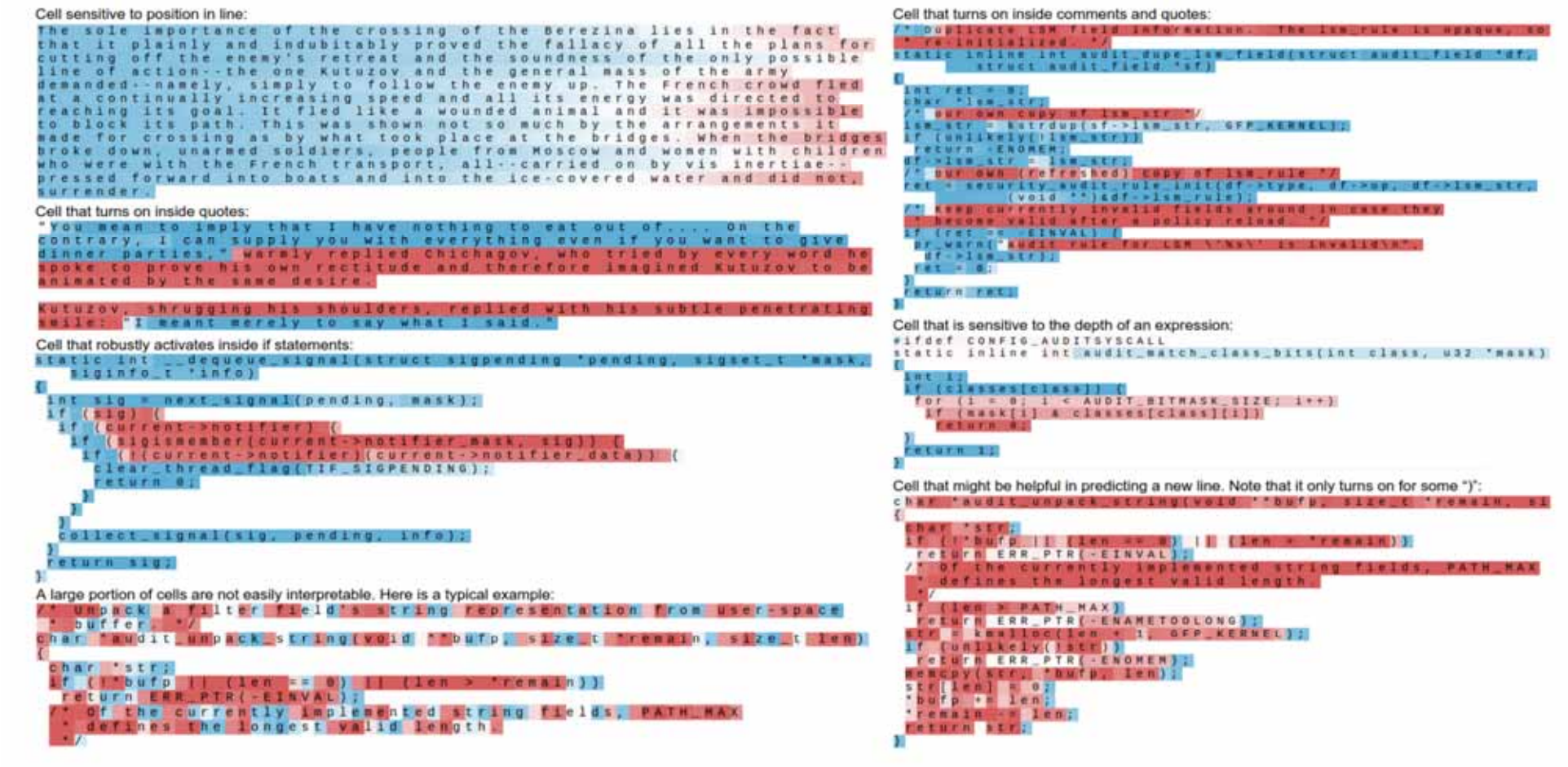
https://www.youtube.com/watch?v=8yI4ubnz-o4

Why does this matter?

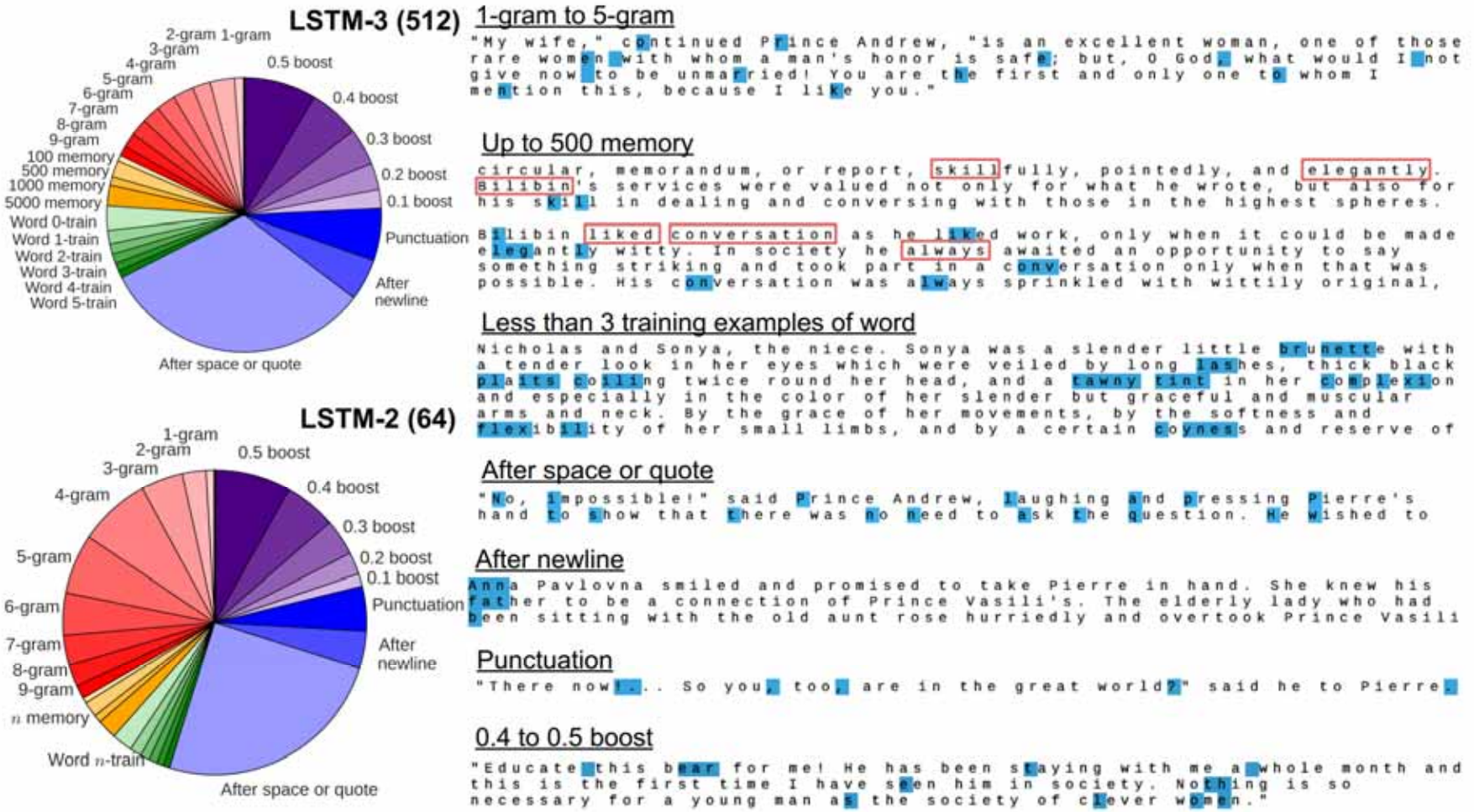
- Interacting with DNNs can teach us about how they work. These interactions can help build our intuitions, which can in turn help us design better models, e.g.:
- Important features such as face detectors and text detectors are learned, even though we do not ask the networks to specifically learn these things. Instead, it learns them because they help with other tasks (e.g. identifying bowties, which are usually paired with faces, and bookcases, which are usually full of books labeled with text, …).
- Common knowledge told that DNN representations are highly distributed, and thus any individual neuron or dimension is uninterpretable (per se). Visualizations by Yosinski et al. suggest that many neurons represent abstract features (e.g. faces, wheels, text, etc.) in a more local, interpretable, manner

Jason Yosinski, Jeff Clune, Anh Nguyen, Thomas Fuchs & Hod Lipson 2015. Understanding neural networks through deep visualization. *arXiv:1506.06579*.

03 Recursive Neural Networks cell state analysis

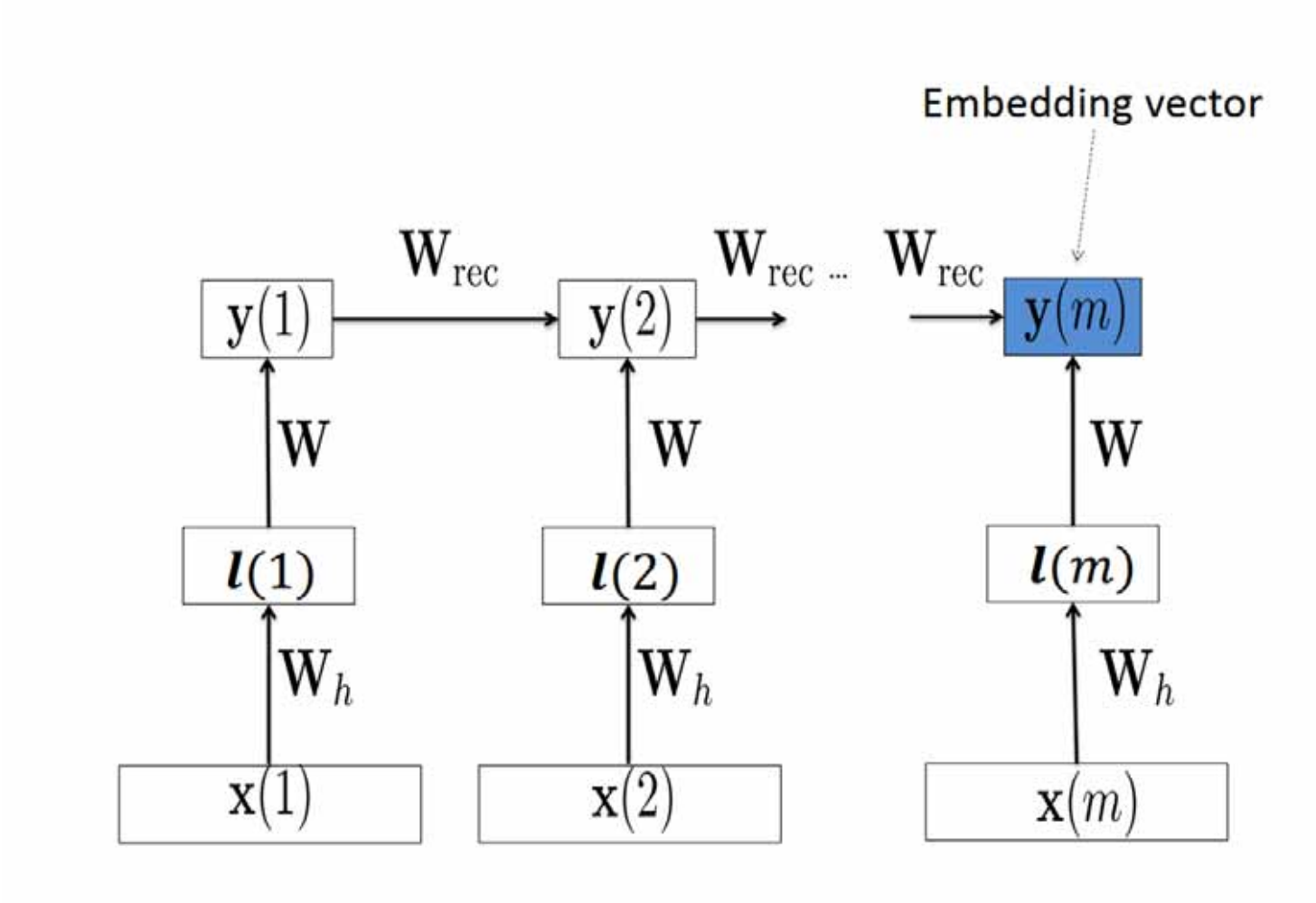


Andrej Karpathy, Justin Johnson & Li Fei-Fei 2015. Visualizing and understanding recurrent networks. arXiv:1506.02078.



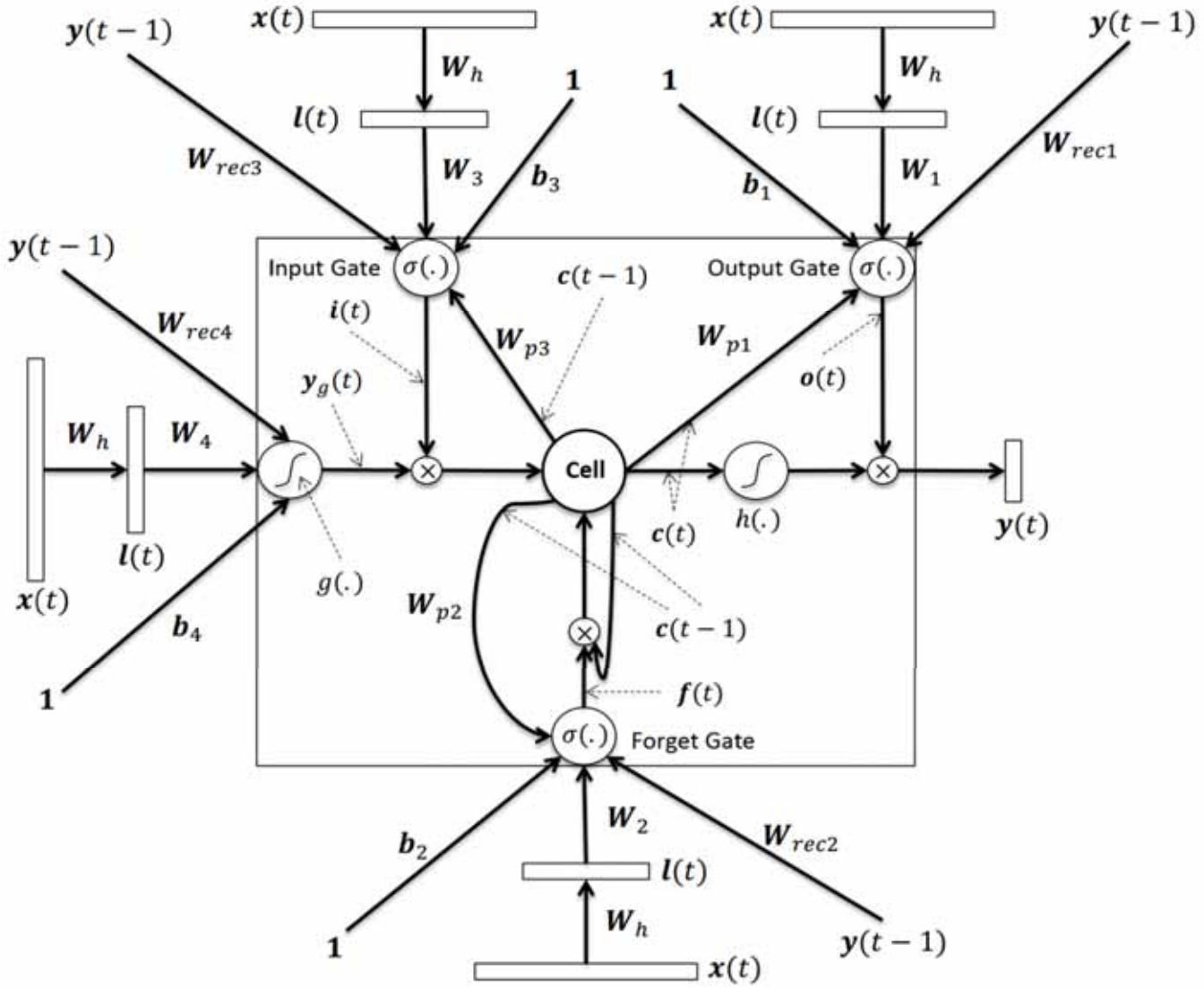
Andrej Karpathy, Justin Johnson & Li Fei-Fei 2015. Visualizing and understanding recurrent networks. arXiv:1506.02078.

Example: RNN for sentence embedding



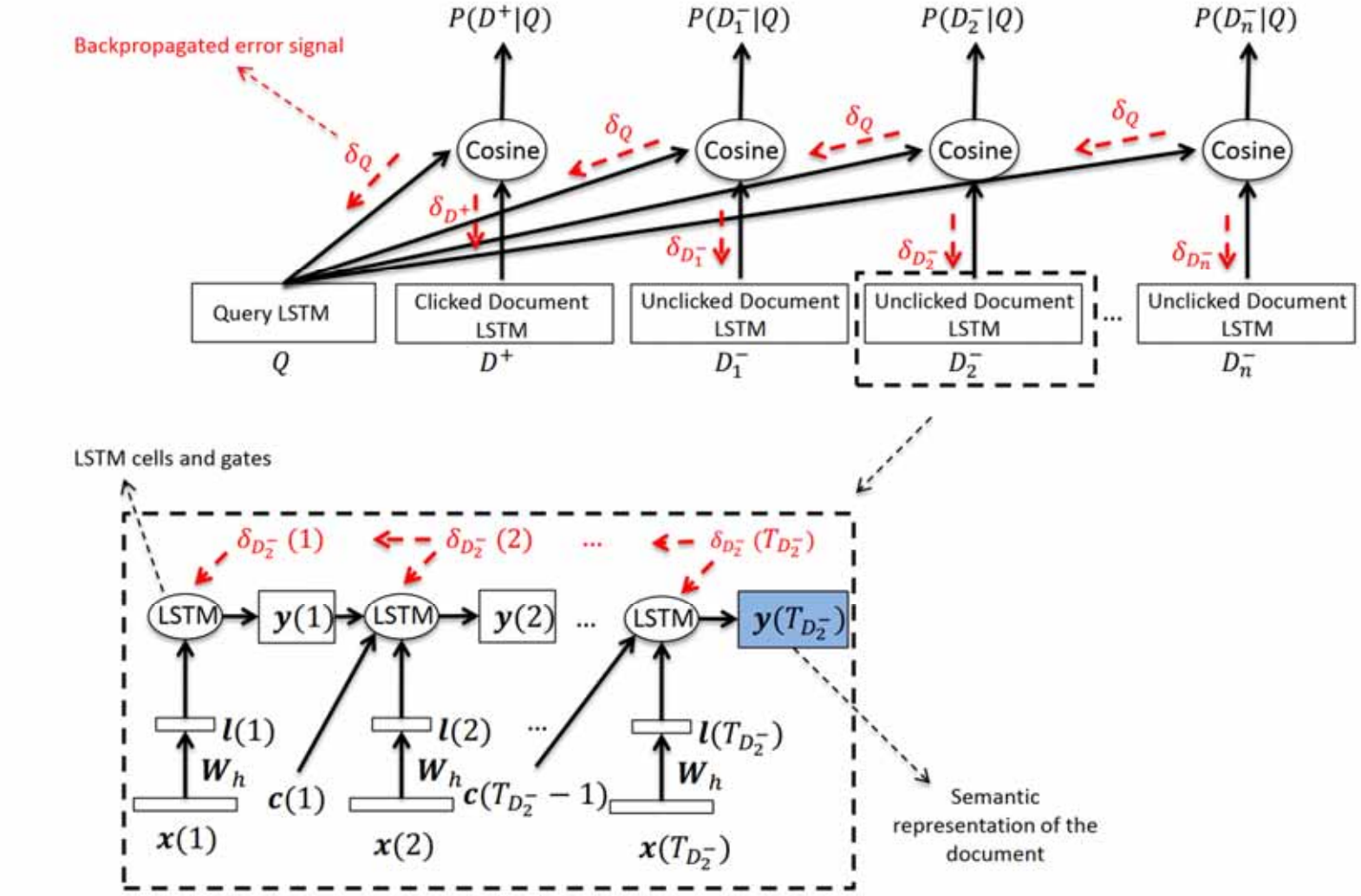
Hamid Palangi, Li Deng, Yelong Shen, Jianfeng Gao, Xiaodong He, Jianshu Chen, Xinying Song & Rabab Ward 2016. Deep sentence embedding using long short-term memory networks: Analysis and application to information retrieval. IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP), 24, (4), 694-707.

Example LSTM for sentence embedding



Hamid Palangi, Li Deng, Yelong Shen, Jianfeng Gao, Xiaodong He, Jianshu Chen, Xinying Song & Rabab Ward 2016. Deep sentence embedding using long short-term memory networks: Analysis and application to information retrieval. IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP), 24, (4), 694-707.

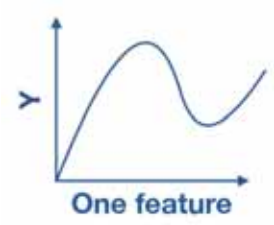
Method by Palangi et al. (2016)



Hamid Palangi, Li Deng, Yelong Shen, Jianfeng Gao, Xiaodong He, Jianshu Chen, Xinying Song & Rabab Ward 2016. Deep sentence embedding using long short-term memory networks: Analysis and application to information retrieval. IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP), 24, (4), 694-707.

04 Fitted Additive

Generalized Additive Models (GAMs)

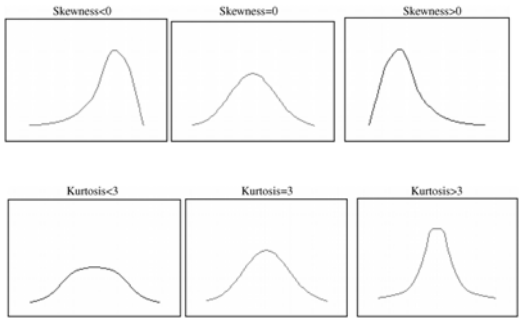


$$y = \beta_0 + \beta_1 x_1 + ... + \beta_n x_n$$

$$g(y) = \beta_0 + \beta_1 x_1 + ... + \beta_n x_n$$

$$g(y) = f_1(x_1) + ... + f_n(x_n)$$

$$g(y) = f_1(x_1) + ... + f_n(x_n) + \sum_{i \neq j} f_{ij}(x_i, x_j).$$



Stéphane Tufféry 2011. Overview of Data Mining. Data Mining and Statistics for Decision Making (Wiley Series in Computational Statistics). New York: John Wiley & Sons, Ltd, pp. 1-24, doi:10.1002/9780470979174.ch1.

$$g(E[y]) = \sum f_i(x_i)$$

Trevor Hastie & Robert Tibshirani 1986. Generalized Additive Models. Statistical Science, 1, (3), 297-318.

Trevor Hastie & Robert Tibshirani 1995. Generalized additive models for medical research. Statistical methods in medical research, 4, (3), 187-196, doi:10.1177/096228029500400302.

Trevor J Hastie 2017. Generalized additive models. Statistical models in S. Routledge, pp. 249-307.

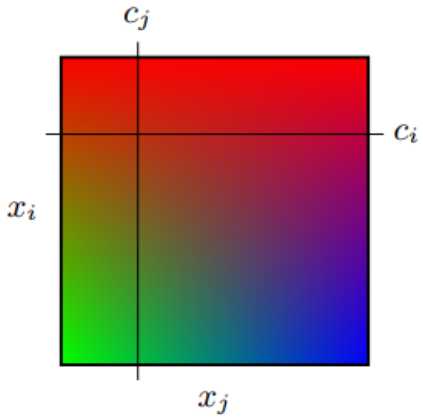
GA²M approach

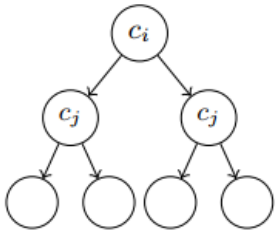
Algorithm 1 GA²M Framework

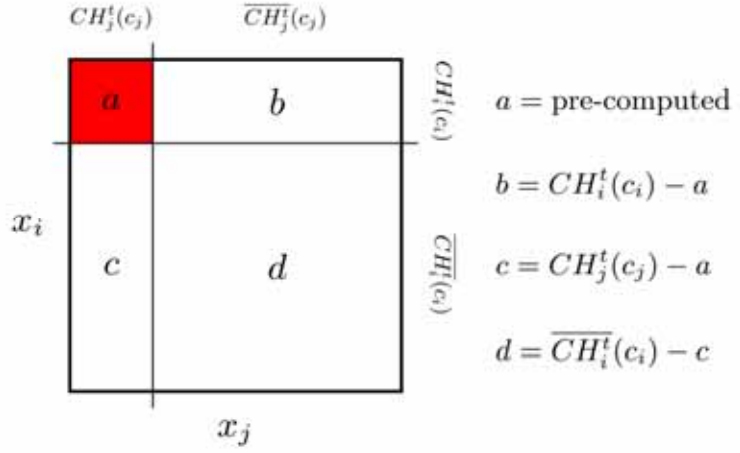
$\mathcal{S} \leftarrow \varnothing$
$\mathcal{Z} \leftarrow \mathcal{U}^2$
while not converge **do**
  $F \leftarrow \arg\min_{F \in \mathcal{H}^1 + \sum_{u \in \mathcal{S}} \mathcal{H}_u} \frac{1}{2} E[(y - F(\boldsymbol{x}))^2]$
  $R \leftarrow y - F(\boldsymbol{x})$
  for all $u \in \mathcal{Z}$ **do**
    $F_u \leftarrow E[R|x_u]$
  $u^* \leftarrow \arg\min_{u \in \mathcal{Z}} \frac{1}{2} E[(R - F_u(x_u))^2]$
  $\mathcal{S} \leftarrow \mathcal{S} \cup \{u^*\}$
  $\mathcal{Z} \leftarrow \mathcal{Z} - \{u^*\}$

Algorithm 2 ConstructLookupTable

$sum \leftarrow 0$
for $q = 1$ to d_j **do**
  $sum \leftarrow sum + H_{ij}^t(v_i^1, v_j^q)$
  $a[1][q] \leftarrow sum$
  $L(v_i^1, v_j^q) \leftarrow ComputeValues(CH_i^t, CH_j^t, a[1][q])$
for $p = 2$ to d_i **do**
  $sum \leftarrow 0$
  for $q = 1$ to d_j **do**
    $sum \leftarrow sum + H_{ij}^t(v_i^p, v_j^q)$
    $a[p][q] \leftarrow sum + a[p-1][q]$
    $L(v_i^p, v_j^q) \leftarrow ComputeValues(CH_i^t, CH_j^t, a[p][q])$

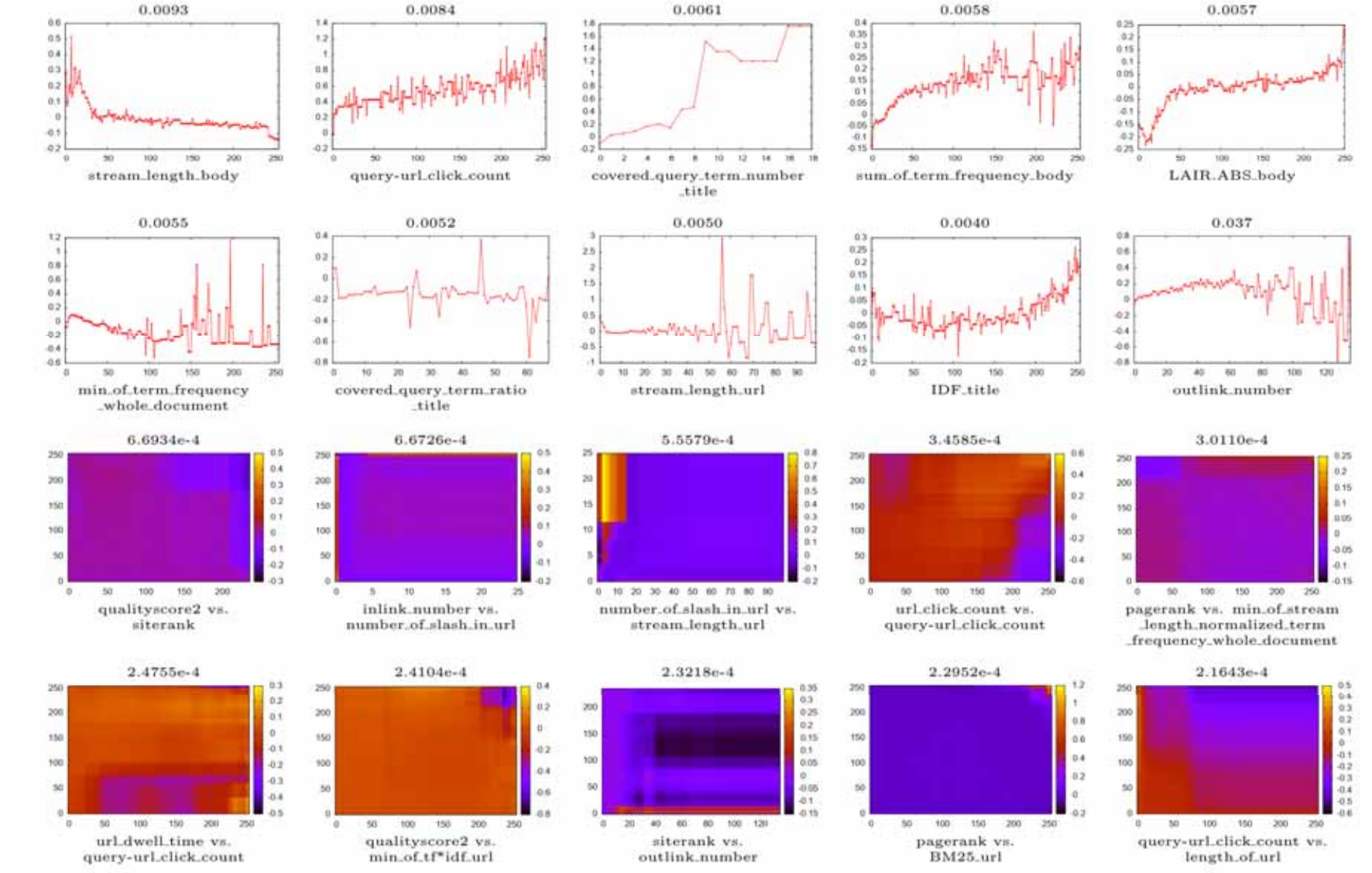






Yin Lou, Rich Caruana, Johannes Gehrke & Giles Hooker. Accurate intelligible models with pairwise interactions. Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining, 2013. ACM, 623-631.

GA^{2}M approach



Yin Lou, Rich Caruana, Johannes Gehrke & Giles Hooker. Accurate intelligible models with pairwise interactions. Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining, 2013. ACM, 623-631.

05 Interactive Machine Learning with the human-in-the-loop

“Solve intelligence – then solve everything else”

Demis Hassabis, 22 May 2015

The Royal Society,
Future Directions of Machine Learning Part 2

Google DeepMind

https://youtu.be/XAbLn66iHcQ?t=1h28m54s

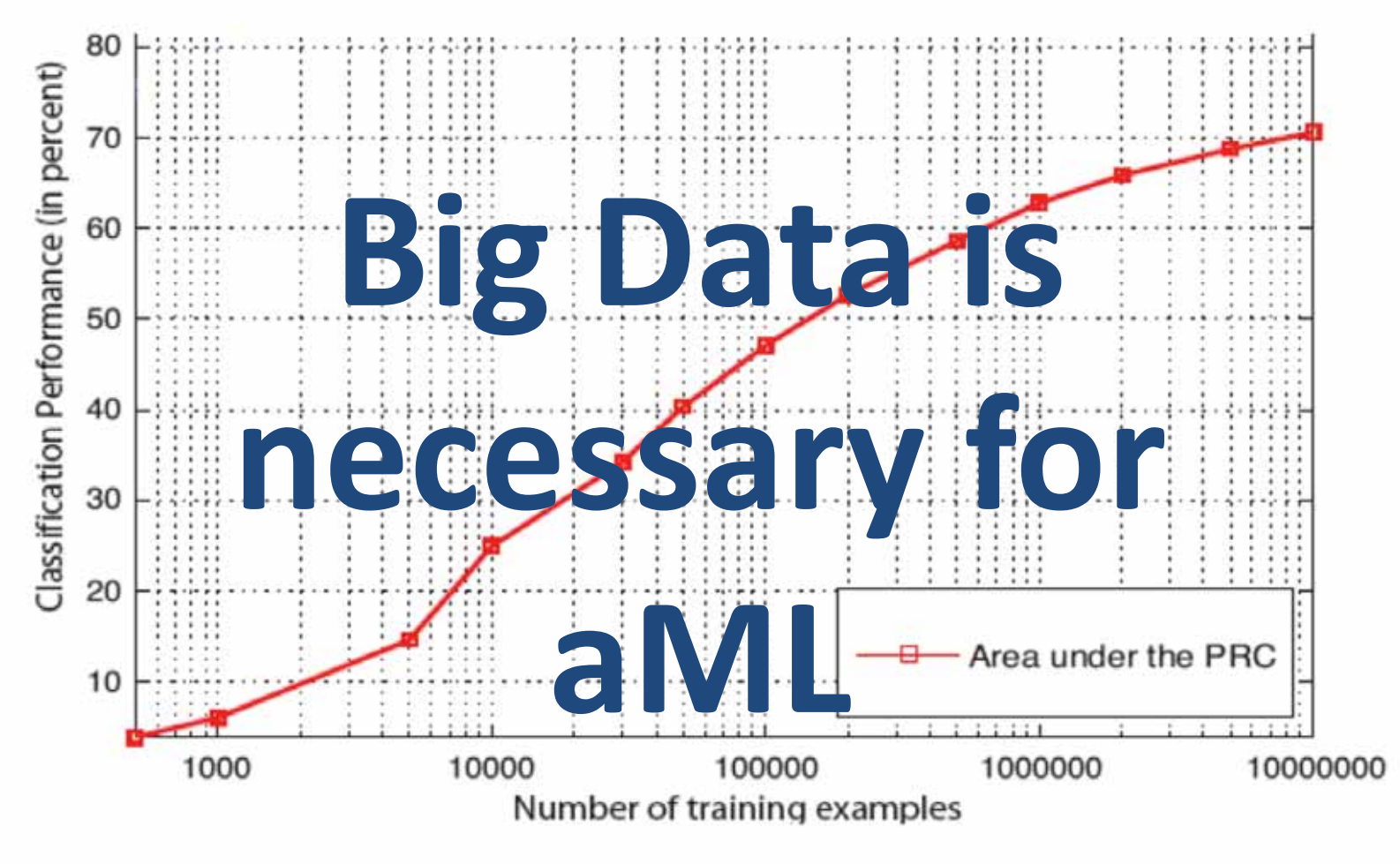
Understanding Intelligence ?

To reach a level of usable intelligence we need to

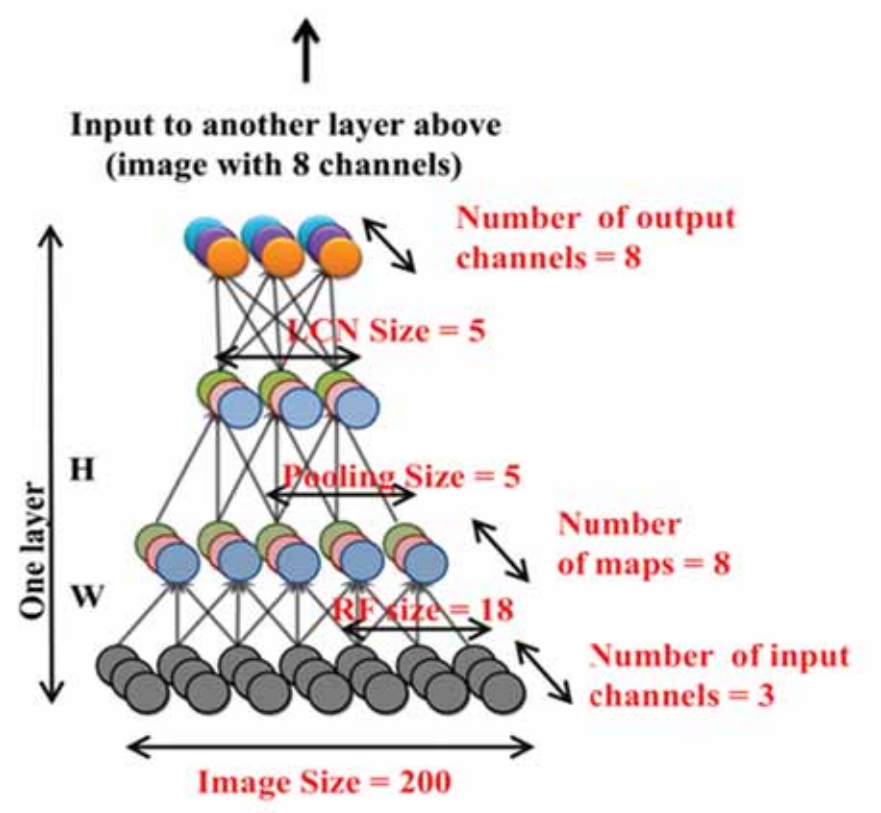
- 1) … learn from prior data
- 2) … extract knowledge
- 2) … generalize, i.e. guessing where a probability mass function concentrates
- 4) … fight the curse of dimensionality
- 5) … disentangle **underlying explanatory factors of data**, i.e.
- 6) … **understand** the data in the **context** of an application domain

Understanding Context !

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.



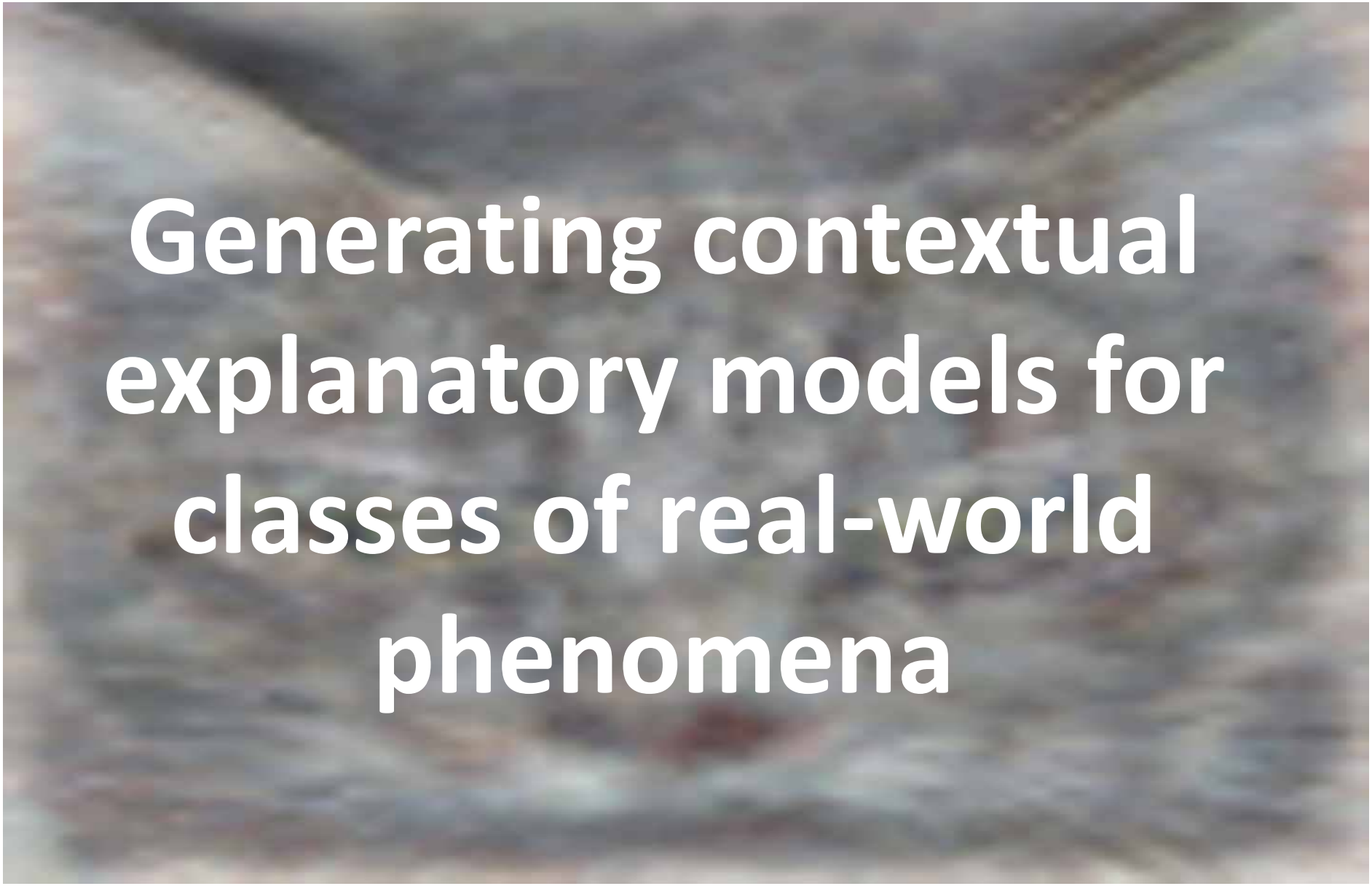
$$x^* = \arg\min_x f(x; W, H), \text{ subject to } ||x||_2 = 1.$$

Le, Q. V., Ranzato, M. A., Monga, R., Devin, M., Chen, K., Corrado, G. S., Dean, J. & Ng, A. Y. 2011. Building high-level features using large scale unsupervised learning. arXiv preprint arXiv:1112.6209.

Le, Q. V. 2013. Building high-level features using large scale unsupervised learning. *IEEE Intl. Conference on Acoustics, Speech and Signal Processing ICASSP.* IEEE. 8595-8598, doi:10.1109/ICASSP.2013.6639343.

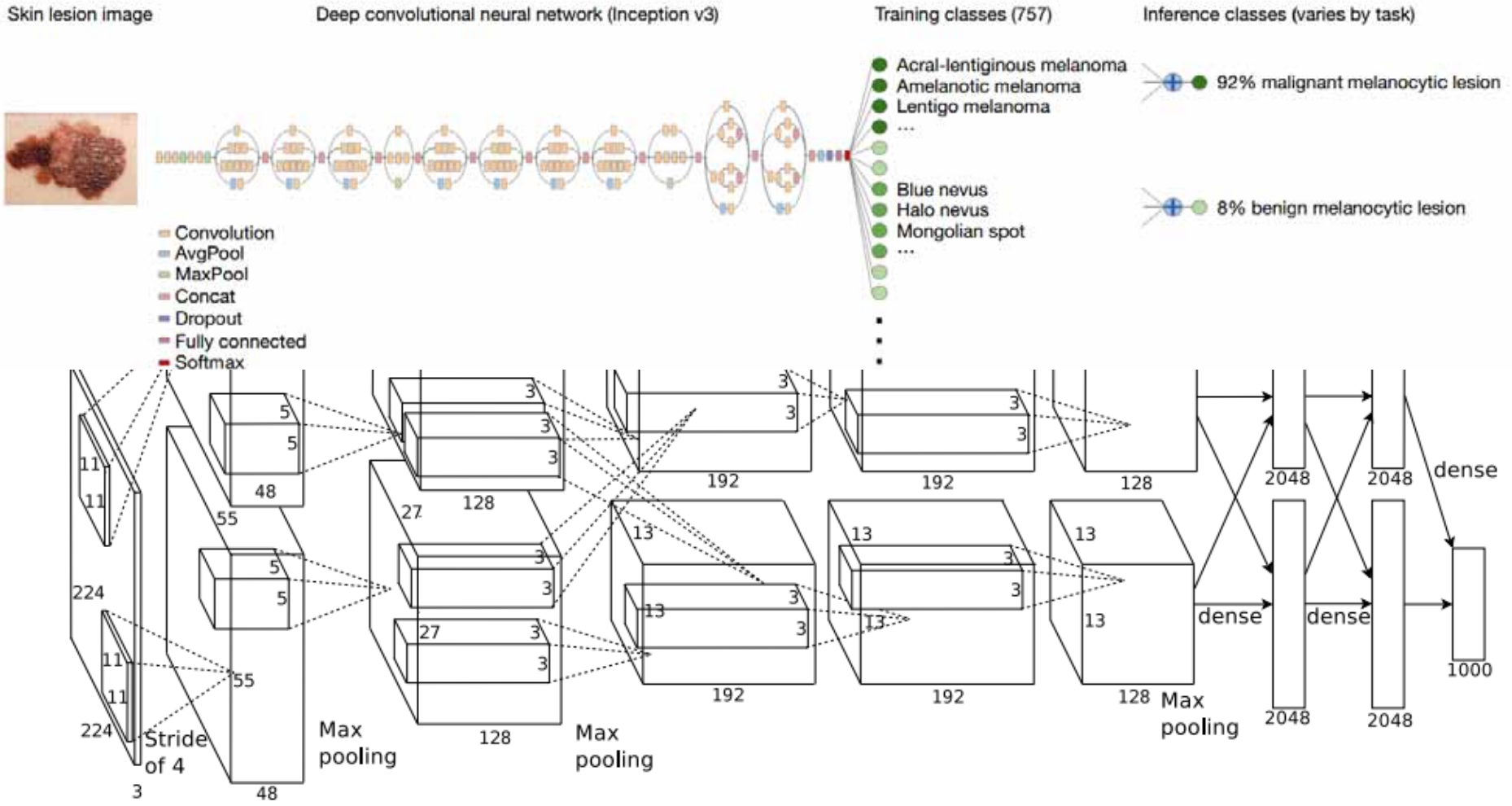
Why is it a cat? What makes a cat a cat?

Image in the public domain, Credit to Kevin Dooley


Generating contextual explanatory models for classes of real-world phenomena

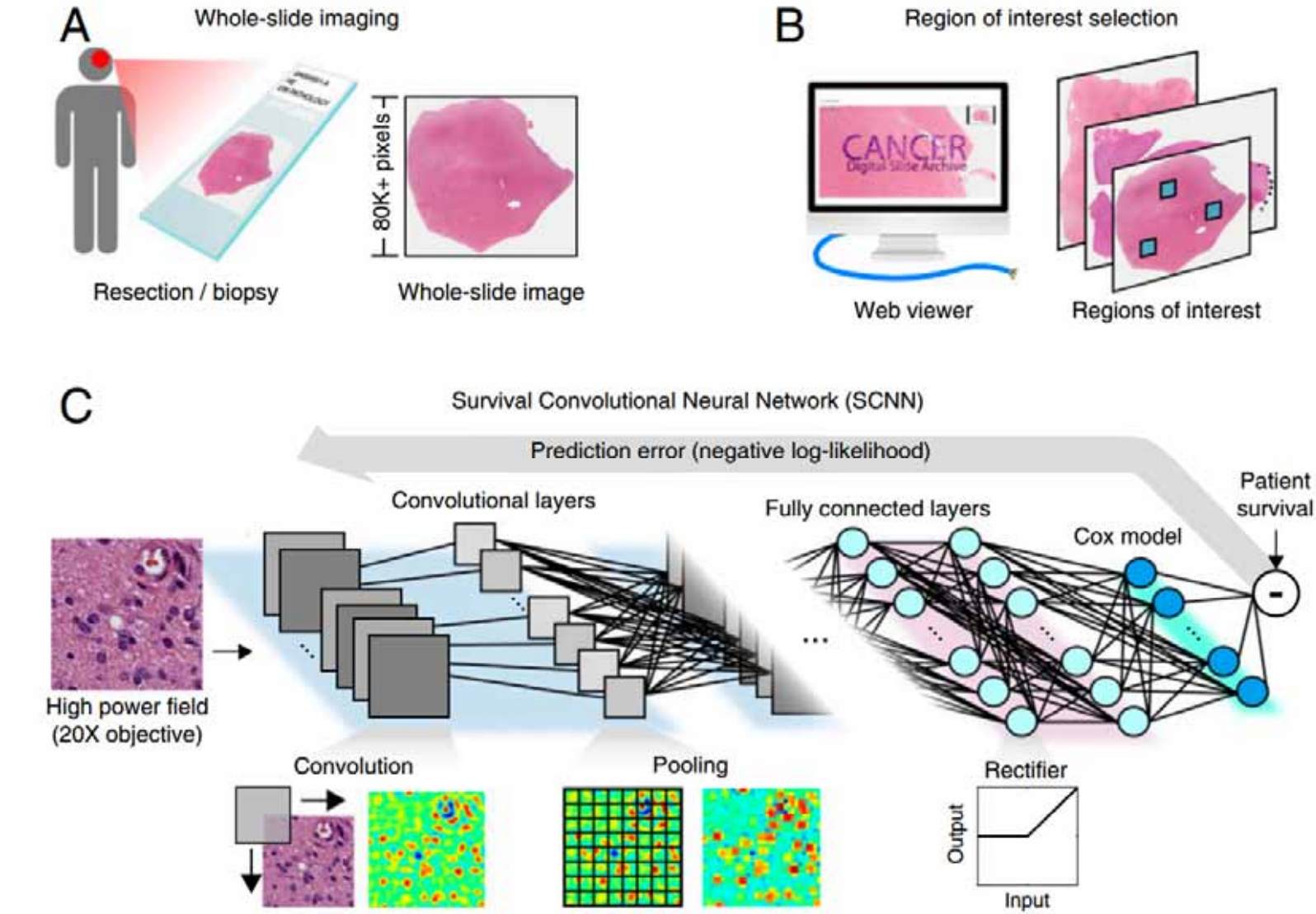
Another famous example: 1,28 million images …

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M. & Thrun, S. 2017. Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542, (7639), 115-118, doi:10.1038/nature21056.



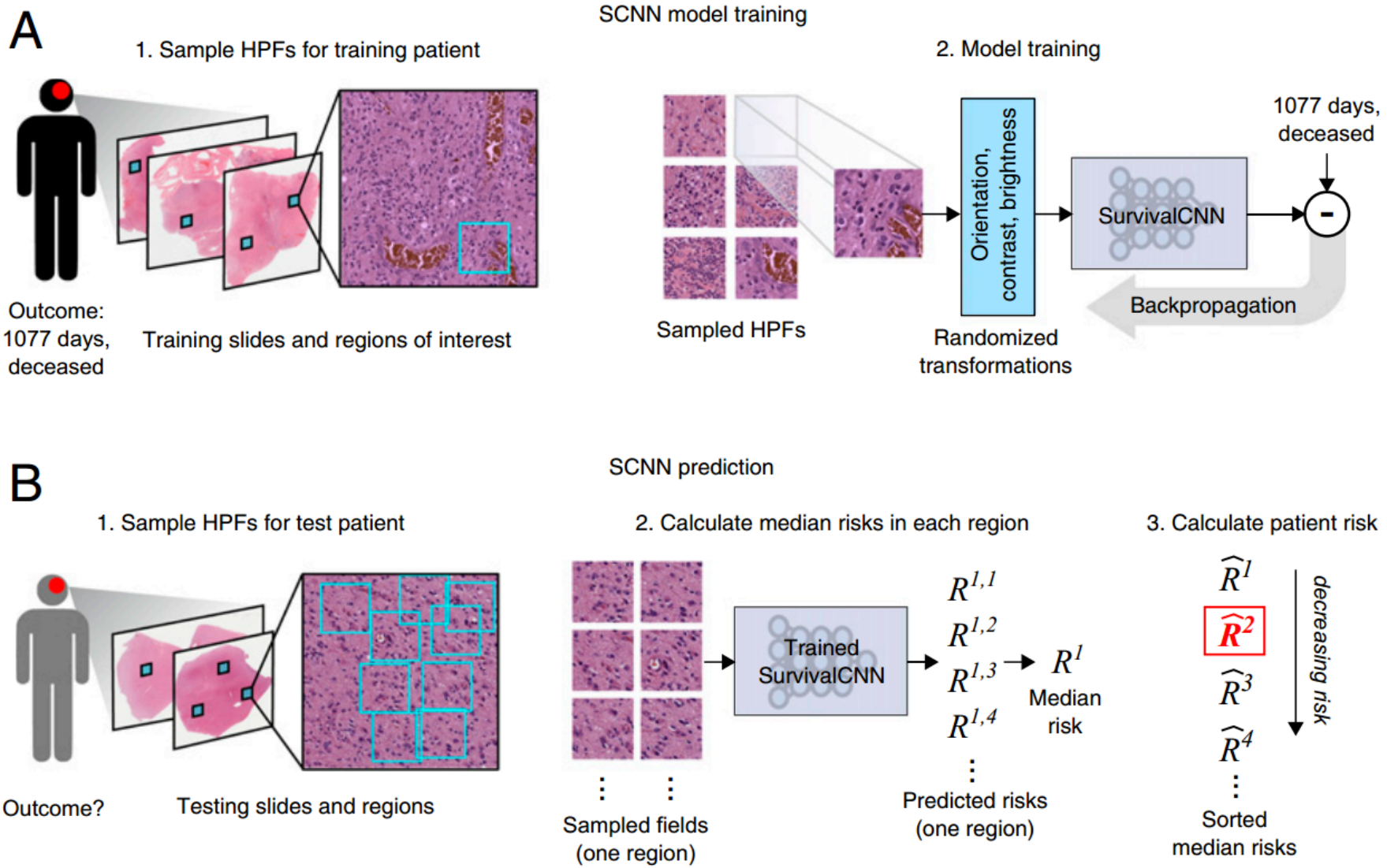
Krizhevsky, A., Sutskever, I. & Hinton, G. E. Imagenet classification with deep convolutional neural networks. In: Pereira, F., Burges, C. J. C., Bottou, L. & Weinberger, K. Q., eds. Advances in neural information processing systems (NIPS 2012), 2012 Lake Tahoe. 1097-1105.

Example from Digital Pathology



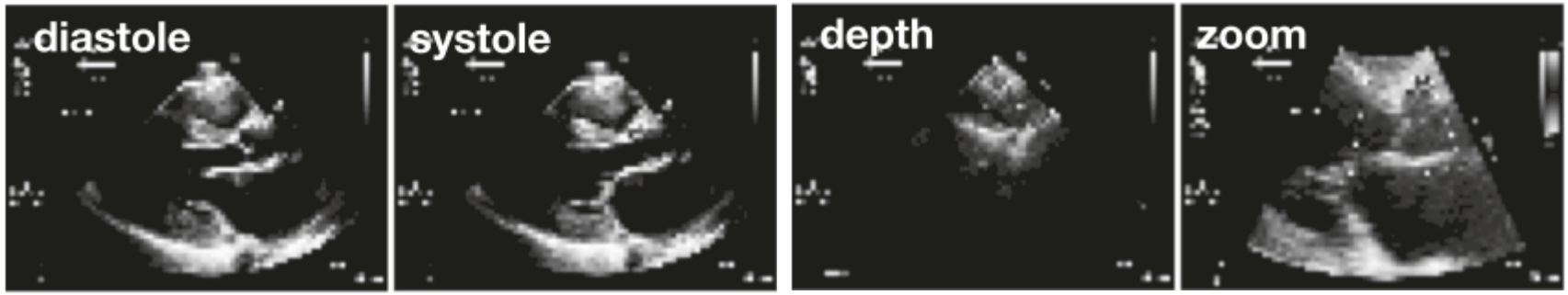
Mobadersany, P., Yousefi, S., Amgad, M., Gutman, D. A., Barnholtz-Sloan, J. S., Velázquez Vega, J. E., Brat, D. J., & Cooper, L. A. D. (2018). Predicting cancer outcomes from histology and genomics using convolutional networks. Proceedings of the National Academy of Sciences. doi: 10.1073/pnas.1717139115

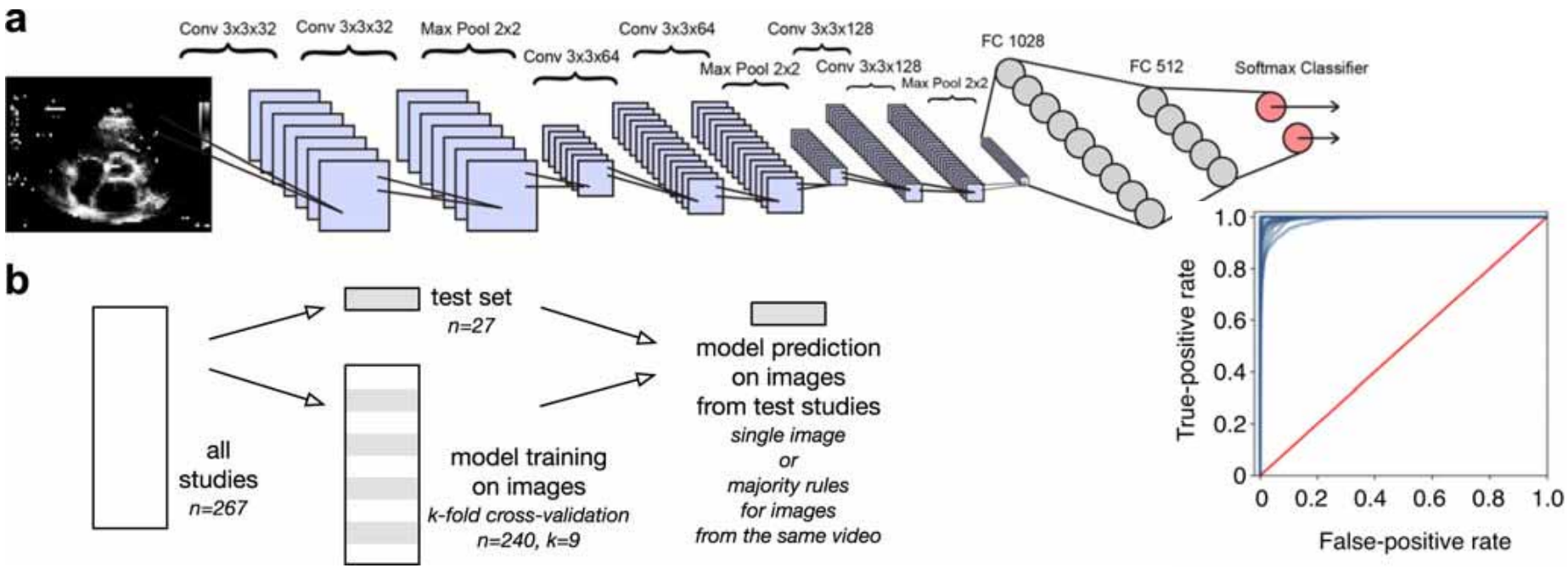
Large amounts of big complex data



Mobadersany, P., Yousefi, S., Amgad, M., Gutman, D. A., Barnholtz-Sloan, J. S., Velázquez Vega, J. E., Brat, D. J., & Cooper, L. A. D. (2018). Predicting cancer outcomes from histology and genomics using convolutional networks. Proceedings of the National Academy of Sciences. doi: 10.1073/pnas.1717139115

Variation in input data needs data augmentation





Madani, A., Arnaout, R., Mofrad, M., & Arnaout, R. (2018). Fast and accurate view classification of echocardiograms using deep learning. Nature Digital Medicine, 1(1), 6. doi: 10.1038/s41746-017-0013-1

- 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,
    - k-Anonymization,
    - Protein-Folding, …

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

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

iML

- iML := algorithms which interact with agents*) and can optimize their learning behaviour through this interaction

*) where the agents can be human bringing in their intuition and contextual knowledge

Holzinger, A. 2016. Interactive Machine Learning (iML). Informatik Spektrum, 39, (1), 64-68, doi:10.1007/s00287-015-0941-6.

Sometimes we need a doctor-in-the-loop

Image Source: 10 Ways Technology is Changing Healthcare http://newhealthypost.com Posted online on April 22, 2018

A group of experts-in-the-loop

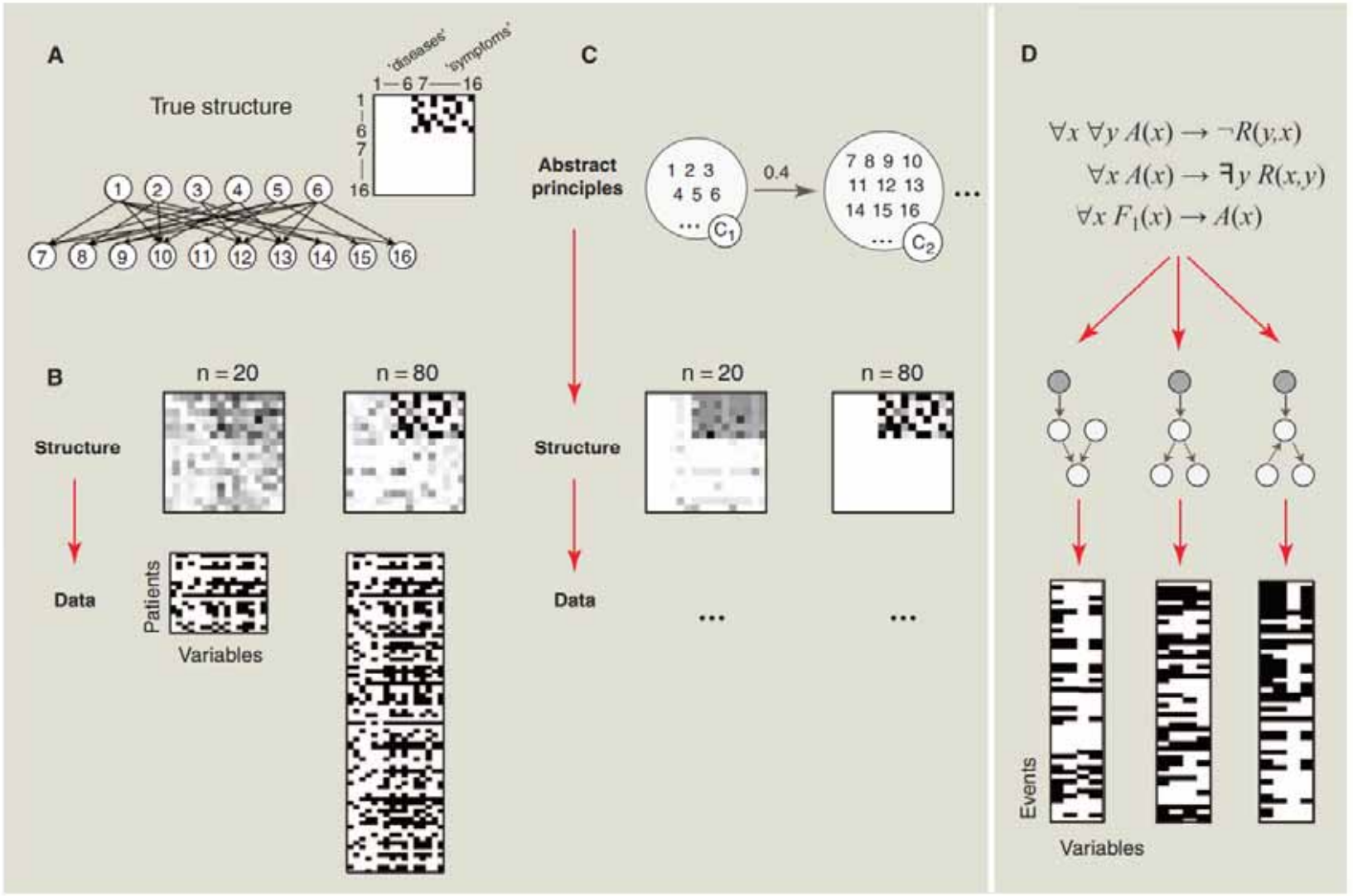
Image Source: Cisco (2008). Cisco Health Presence Trial at Aberdeen Royal Infirmary in Scotland

A crowd of people-in-the-loop

- **Humans can generalize even from few examples …**
  - They learn relevant representations
  - Can disentangle the explanatory factors
  - Find the shared underlying explanatory factors, in particular between $P(x)$ and $P(Y|X)$, with a causal link between $Y \rightarrow X$

Bengio, Y., Courville, A. & Vincent, P. 2013. Representation learning: A review and new perspectives. IEEE transactions on pattern analysis and machine intelligence, 35, (8), 1798-1828, doi:10.1109/TPAMI.2013.50.

Even children can infer from little data



Joshua B. Tenenbaum, Charles Kemp, Thomas L. Griffiths & Noah D. Goodman 2011. How to grow a mind: Statistics, structure, and abstraction. Science, 331, (6022), 1279-1285, doi:10.1126/science.1192788.

Adversarial Examples that Fool both Computer Vision and Time-Limited Humans

Gamaleldin F. Elsayed*
Google Brain
gamaleldin.elsayed@gmail.com

Shreya Shankar
Stanford University

Brian Cheung
UC Berkeley

Nicolas Papernot
Pennsylvania State University

Alex Kurakin
Google Brain

Ian Goodfellow
Google Brain

Jascha Sohl-Dickstein
Google Brain
jaschasd@google.com

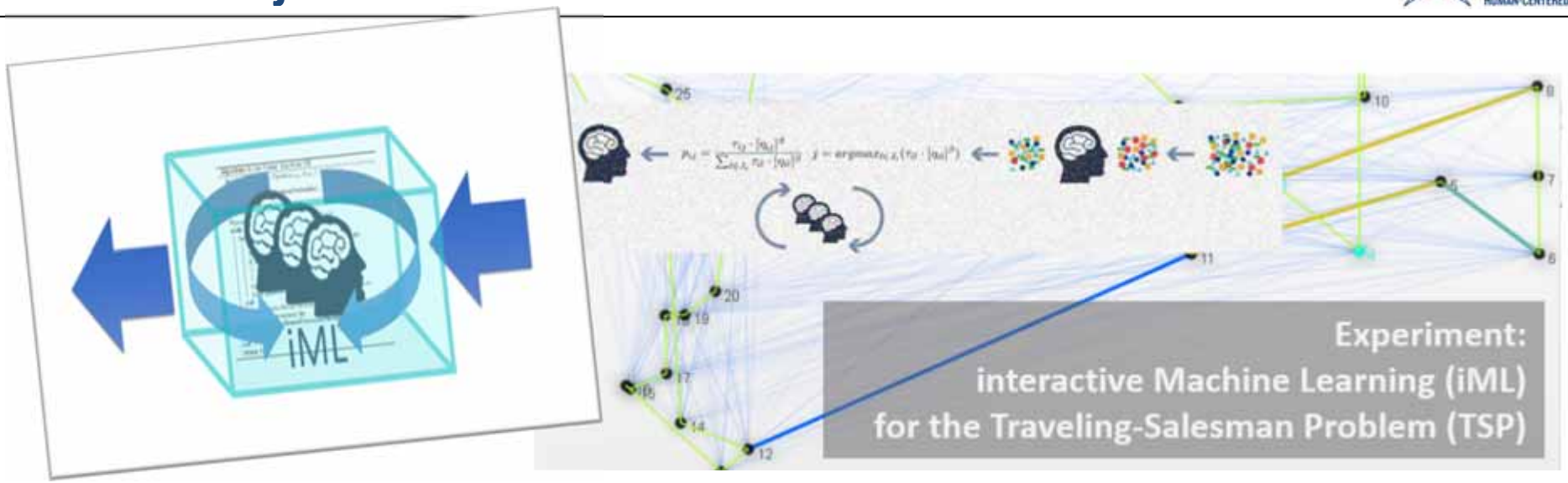
Abstract

Machine learning models are vulnerable to **adversarial examples**: small changes to images can cause computer vision models to make mistakes such as identifying a school bus as an ostrich. However, it is still an open question whether humans are prone to similar mistakes. Here, we address this question by leveraging recent techniques that transfer adversarial examples from computer vision models with known parameters and architecture to other models with unknown parameters and architecture, and by matching the initial processing of the human visual system. We find that adversarial examples that strongly transfer across computer vision models influence the classifications made by time-limited human observers.

See also: Ian J. Goodfellow, Jonathon Shlens & Christian Szegedy 2014. Explaining and harnessing adversarial examples. arXiv:1412.6572.

Gamaleldin F Elsayed, Shreya Shankar, Brian Cheung, Nicolas Papernot, Alex Kurakin, Ian Goodfellow & Jascha Sohl-Dickstein 2018. Adversarial Examples that Fool both Human and Computer Vision. arXiv:1802.08195.

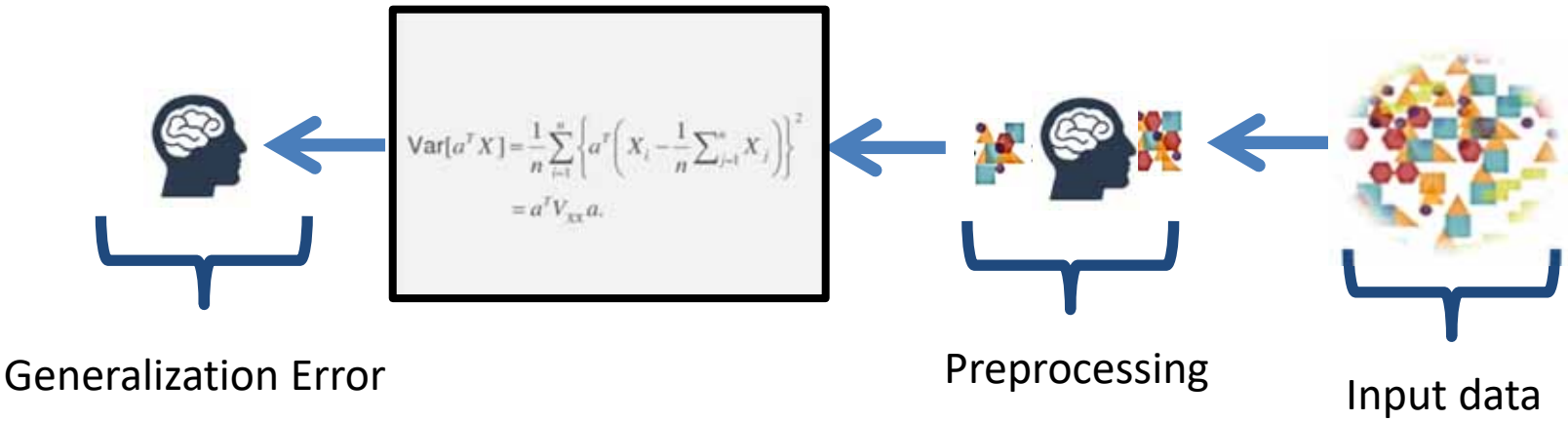
Project: iML

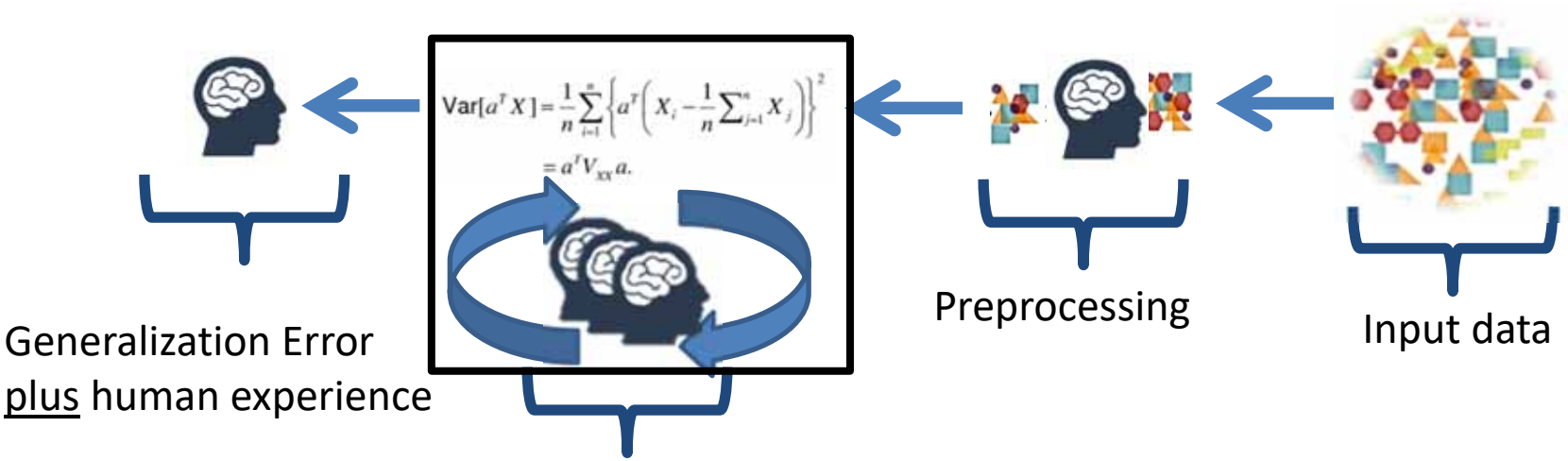


- From black-box to glass-box ML
- Exploit human intelligence for solving hard problems (e.g. Subspace Clustering, k-Anonymization, Protein-Design)
- Towards multi-agent systems with humans-in-the-loop

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. 81-95, doi:10.1007/978-3-319-45507-56.

How can iML help here?

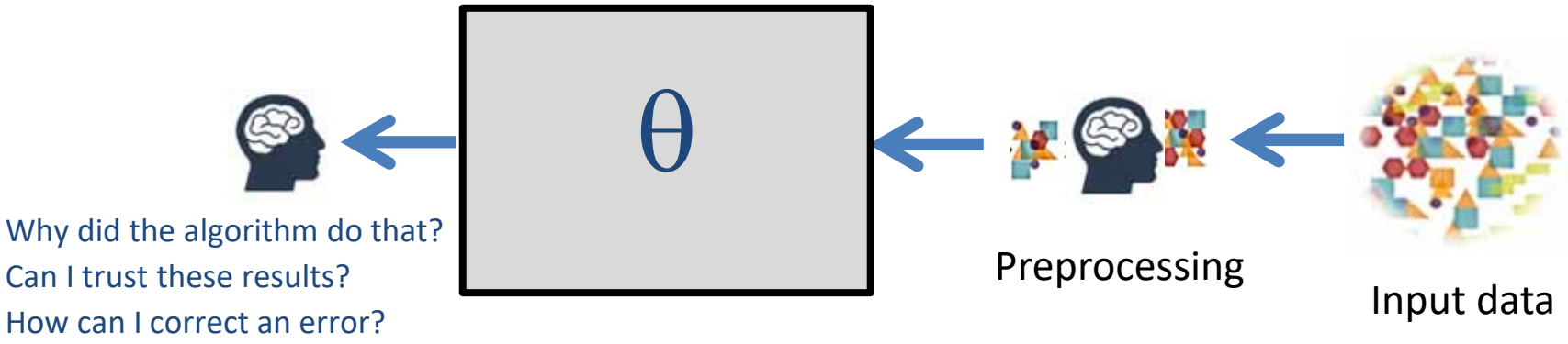


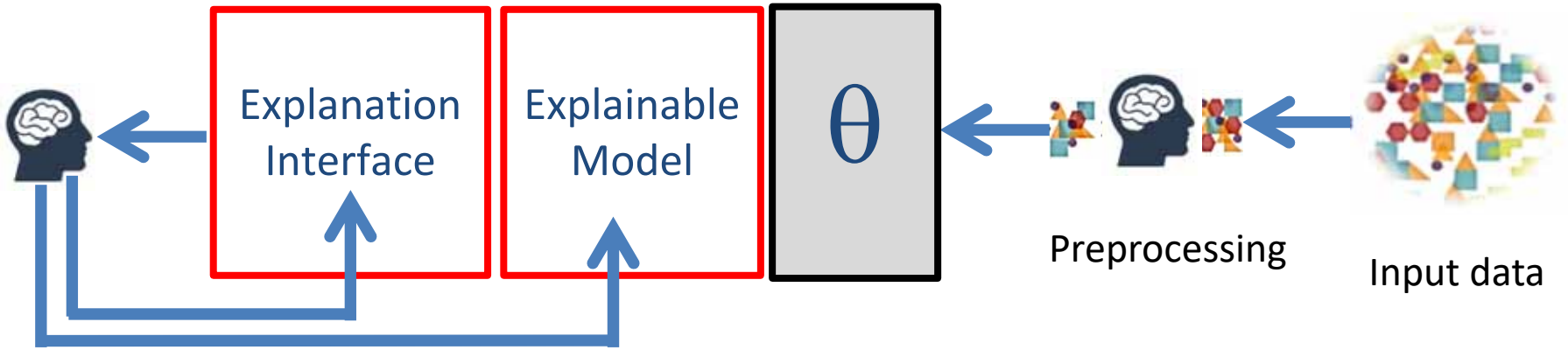


iML = human interaction – bringing in human intuition

Andreas Holzinger, Markus Plass, Katharina Holzinger, Gloria Cerasela Crisan, Camelia-M. Pintea & Vasile Palade 2017. A glass-box interactive machine learning approach for solving NP-hard problems with the human-in-the-loop. arXiv:1708.01104.

Why is this relevant for the medical domain?

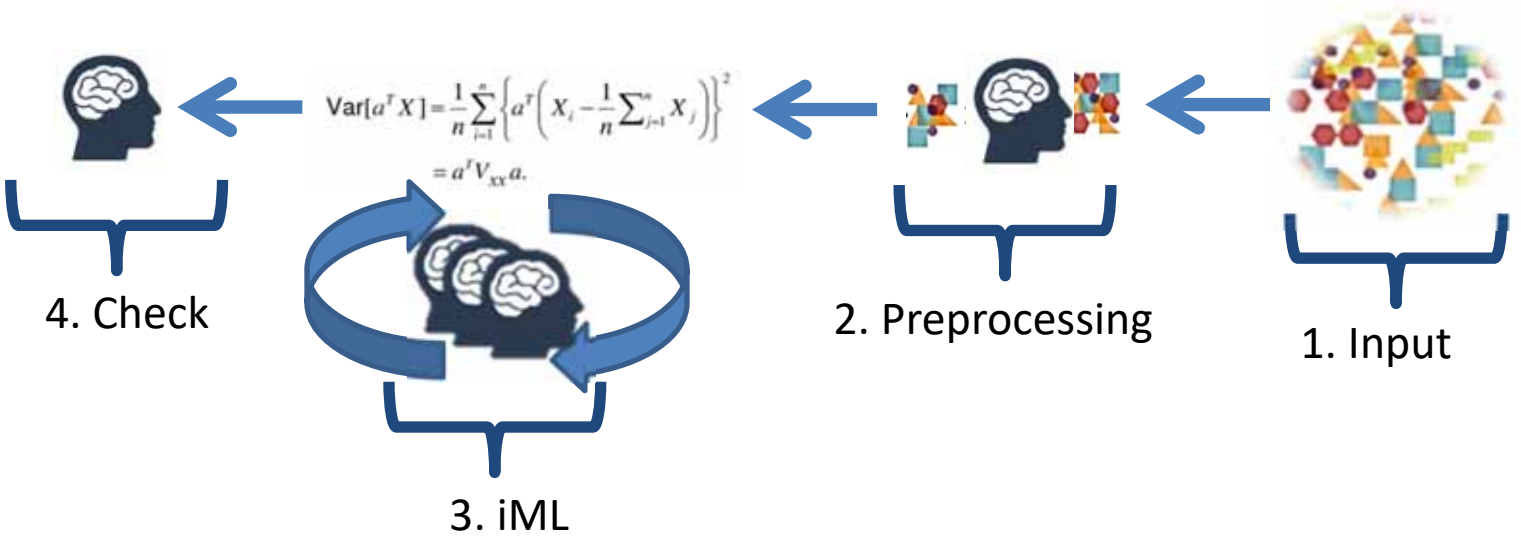




The domain expert can re-enact on demand
The domain expert can understand why a certain machine decision has been made
The domain expert can learn and correct errors

Andreas Holzinger, Chris Biemann, Constantinos S. Pattichis & Douglas B. Kell 2017. What do we need to build explainable AI systems for the medical domain? arXiv:1712.09923.

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 …



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.

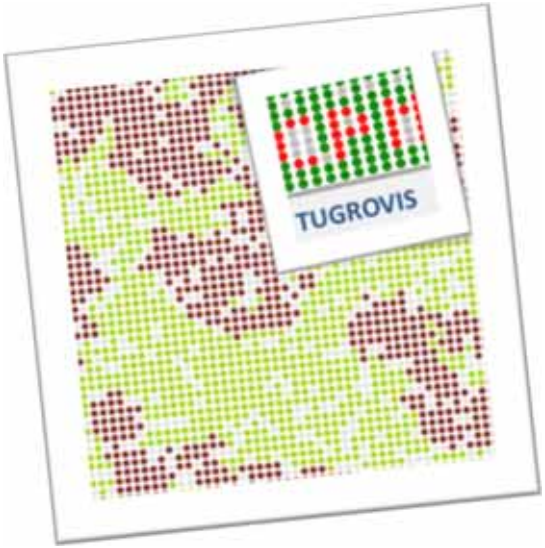
- Example 1: Subspace Clustering
- Example 2: k-Anonymization
- Example 3: Protein Design

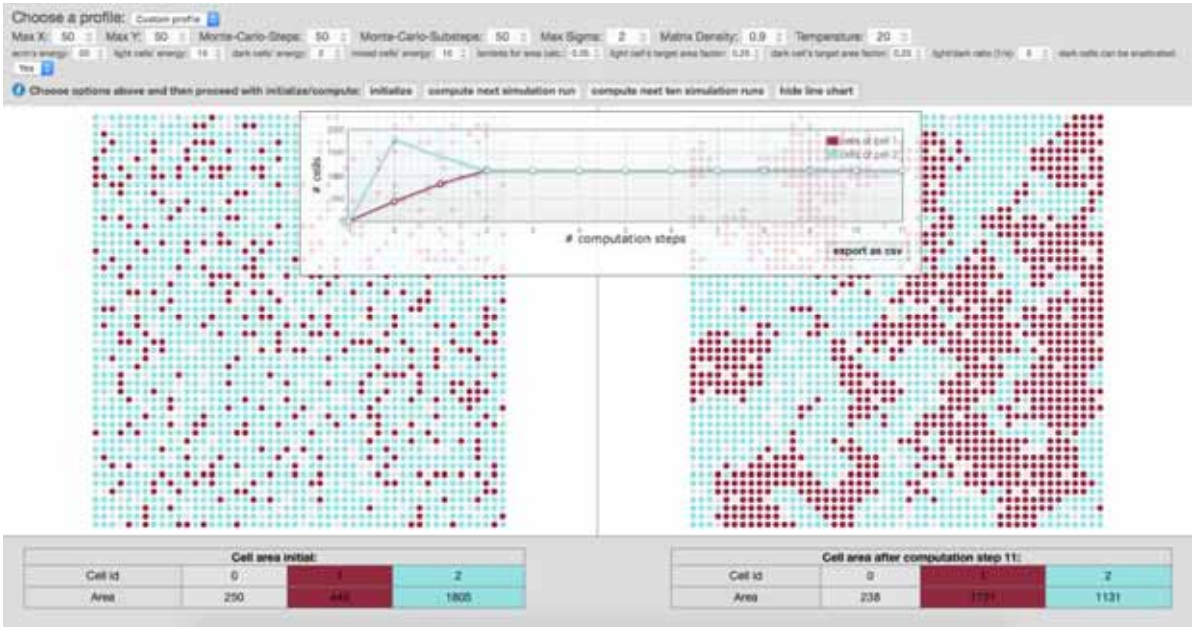
Hund, M., Böhm, D., Sturm, W., Sedlmair, M., Schreck, T., Ullrich, T., Keim, D. A., Majnaric, L. & Holzinger, A. 2016. Visual analytics for concept exploration in subspaces of patient groups: Making sense of complex datasets with the Doctor-in-the-loop. Brain Informatics, 1-15, doi:10.1007/s40708-016-0043-5.

Kieseberg, P., Malle, B., Fruehwirt, P., Weippl, E. & Holzinger, A. 2016. A tamper-proof audit and control system for the doctor in the loop. Brain Informatics, 3, (4), 269–279, doi:10.1007/s40708-016-0046-2.

Lee, S. & Holzinger, A. 2016. Knowledge Discovery from Complex High Dimensional Data. In: Michaelis, S., Piatkowski, N. & Stolpe, M. (eds.) Solving Large Scale Learning Tasks. Challenges and Algorithms, Lecture Notes in Artificial Intelligence LNAI 9580. Springer, pp. 148-167, doi:10.1007/978-3-319-41706-6_7.

Project: Tumor-Growth Simulation



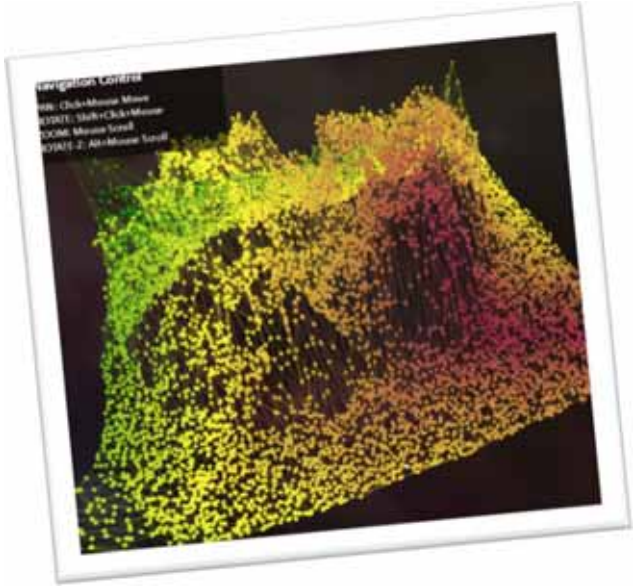


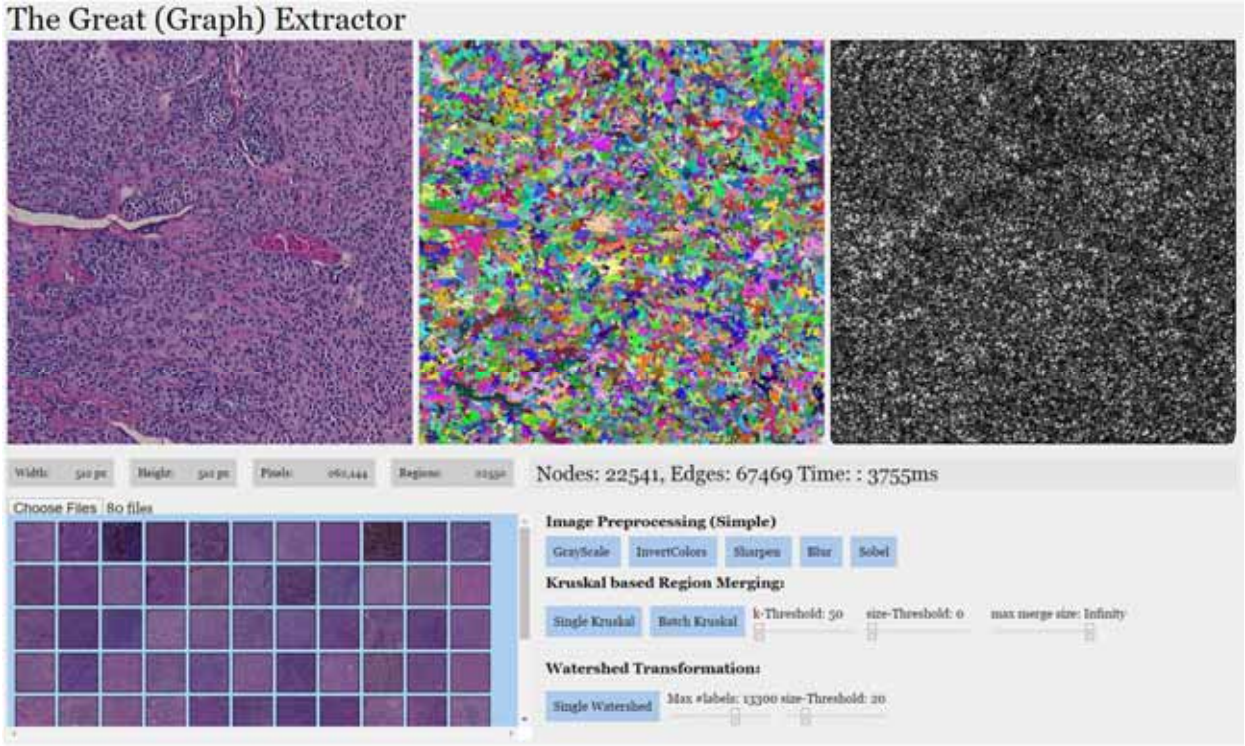
- Contribute to understanding tumor growth
- Goal: Help to Refine → Reduce → Replace
- Towards discrete Multi-Agent Hybrid Systems

Jeanquartier, F., Jean-Quartier, C., Cemernek, D. & Holzinger, A. 2016. In silico modeling for tumor growth visualization. BMC Systems Biology, 10, (1), 1-15, doi:10.1186/s12918-016-0318-8.

Jeanquartier, F., Jean-Quartier, C., Kotlyar, M., Tokar, T., Hauschild, A.-C., Jurisica, I. & Holzinger, A. 2016. Machine Learning for In Silico Modeling of Tumor Growth. In: Springer Lecture Notes in Artificial Intelligence LNAI 9605. Cham: Springer International Publishing, pp. 415-434, doi:10.1007/978-3-319-50478-0_21.

Project: Graphinius

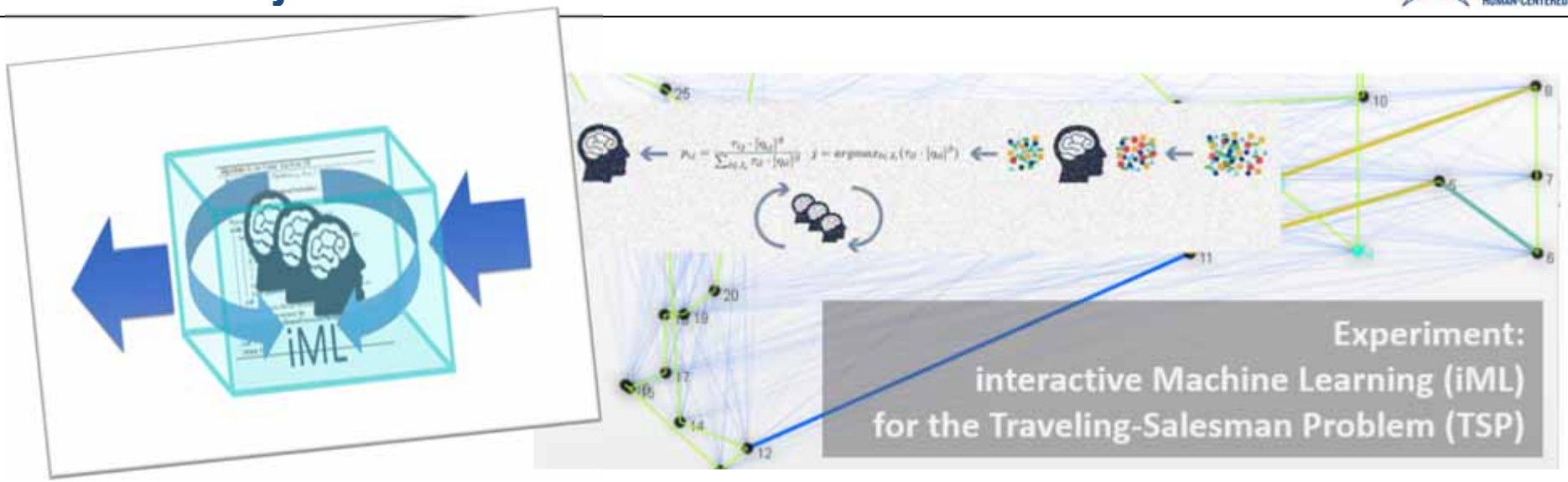




- Contribute to graph understanding and algorithm prototyping by real-time visualization, interaction and manipulation
- Supports client-based federated learning
- Towards an online graph exploration and analysis platform

Malle, B., Kieseberg, P., Weippl, E. & Holzinger, A. 2016. The right to be forgotten: Towards Machine Learning on perturbed knowledge bases. Springer Lecture Notes in Computer Science LNCS 9817. Heidelberg, Berlin, New York: Springer, pp. 251-256, doi:10.1007/978-3-319-45507-5_17.

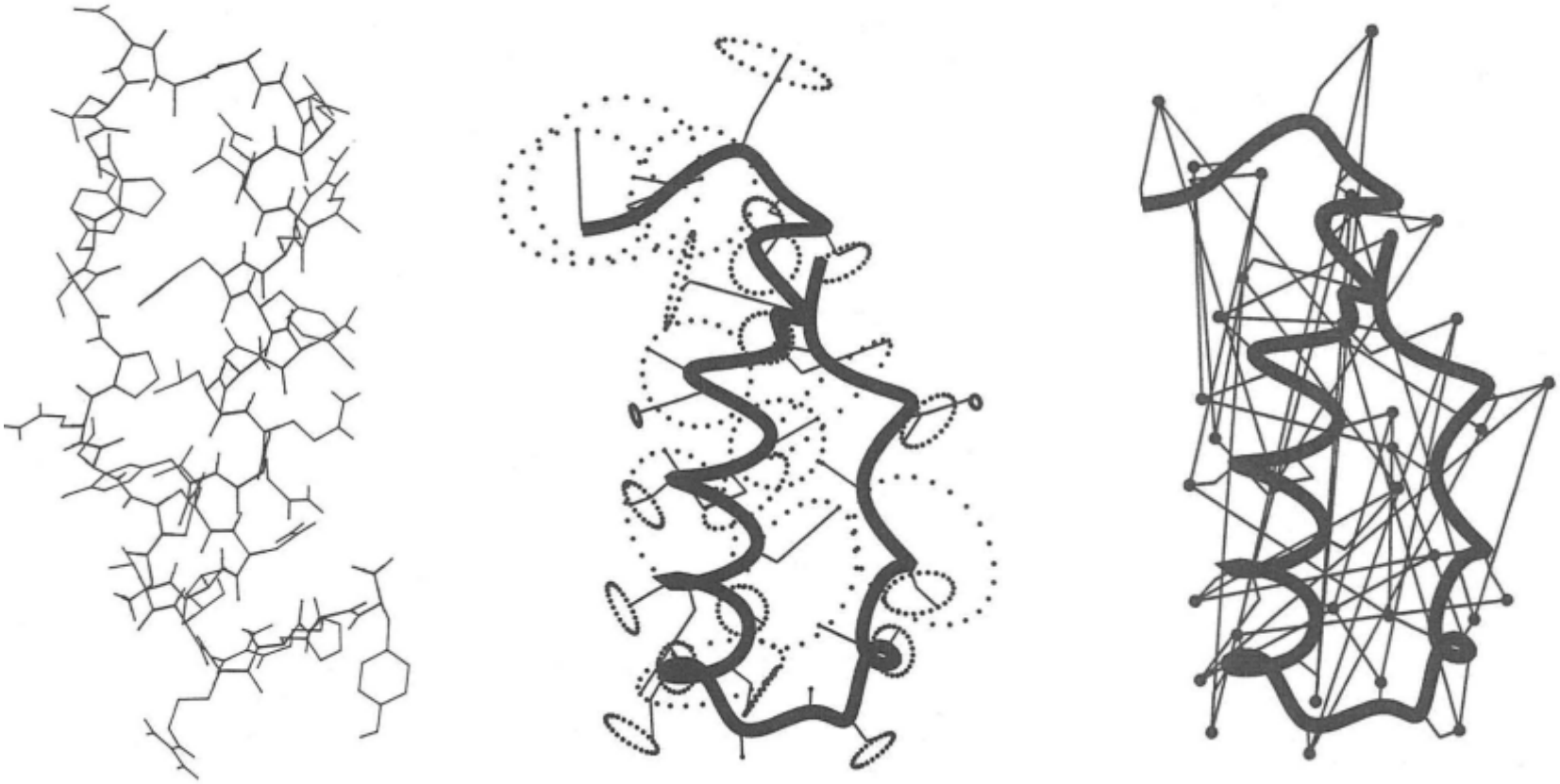
Project: iML



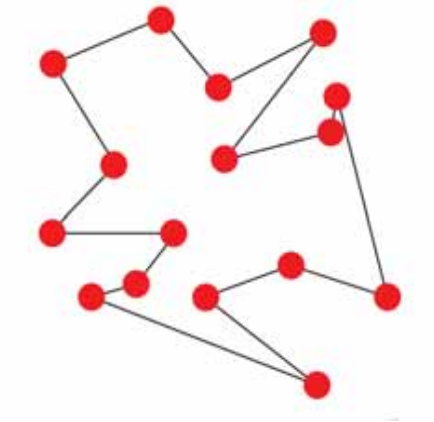
- From black-box to glass-box ML
- Exploit human intelligence for solving hard problems (e.g. Subspace Clustering, k-Anonymization, Protein-Design)
- Towards multi-agent systems with humans-in-the-loop

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. 81-95, doi:10.1007/978-3-319-45507-56.

Protein Folding is a TSP



Bohr, H. & Brunak, S. 1989. A travelling salesman approach to protein conformation. Complex Systems, 3, 9-28



- How many different routes for N=100 ?
- 10^{155}
- We need heuristics and human intuition in the loop of a nature-inspired multi-agent approach

Sim, K. M. & Sun, W. H. 2003. Ant colony optimization for routing and load-balancing: Survey and new directions. IEEE Transactions on Systems Man and Cybernetics Part a-Systems and Humans, 33, (5), 560-572, doi:10.1109/tsmca.2003.817391.

Macgregor, J. N. & Ormerod, T. 1996. Human performance on the traveling salesman problem. Perception & Psychophysics, 58, (4), 527-539, doi:10.3758/bf03213088.

Input : ProblemSize, m, β, ρ, σ, q_0
Output: $Pbest$
$Pbest \leftarrow$ CreateHeuristicSolution(ProblemSize);
$Pbest_{cost} \leftarrow$ Cost($Pbest$);
$Pheromone_{init} \leftarrow \frac{1.0}{ProblemSize \times Pbest_{cost}}$;
$Pheromone \leftarrow$ InitializePheromone($Pheromone_{init}$);
while $\neg StopCondition()$ **do**
  for $i = 1$ to m **do**
    $S_i \leftarrow$ ConstructSolution(Pheromone, ProblemSize, β, q_0);
    $Si_{cost} \leftarrow$ Cost(S_i);
    if $Si_{cost} \leq$ Pbest$_{cost}$ **then**
      $Pbest_{cost} \leftarrow Si_{cost}$;
      $Pbest \leftarrow S_i$;
    end
    LocalUpdateAndDecayPheromone(Pheromone, S_i, Si_{cost}, ρ);
  end
  GlobalUpdateAndDecayPheromone(Pheromone, $Pbest$, $Pbest_{cost}$, ρ);
  while $isUserInteraction()$ **do**
    GlobalAddAndRemovePheromone(Pheromone, $Pbest$, $Pbest_{cost}$, ρ);
  end
end
return P_{best};

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. 81-95, doi:10.1007/978-3-319-45507-56.

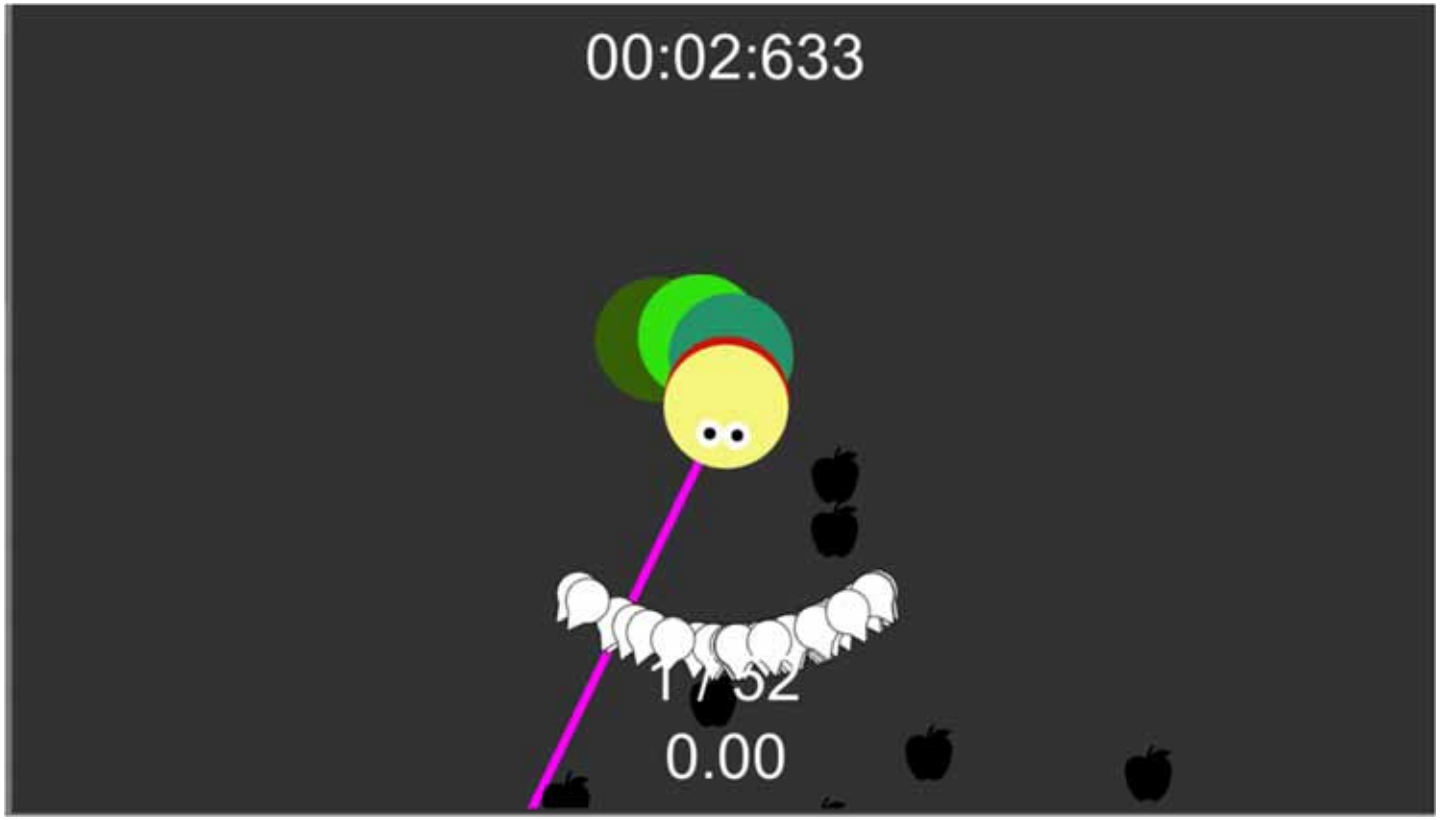
Influence the Pheromone Level by human intuition

$$p_{ij} = \frac{[\tau_{ij}]^{\alpha} \cdot [\eta_{ij}]^{\beta}}{\sum_{l \in J_i^k} [\tau(t)]^{\alpha} \cdot [\eta]^{\beta}}$$

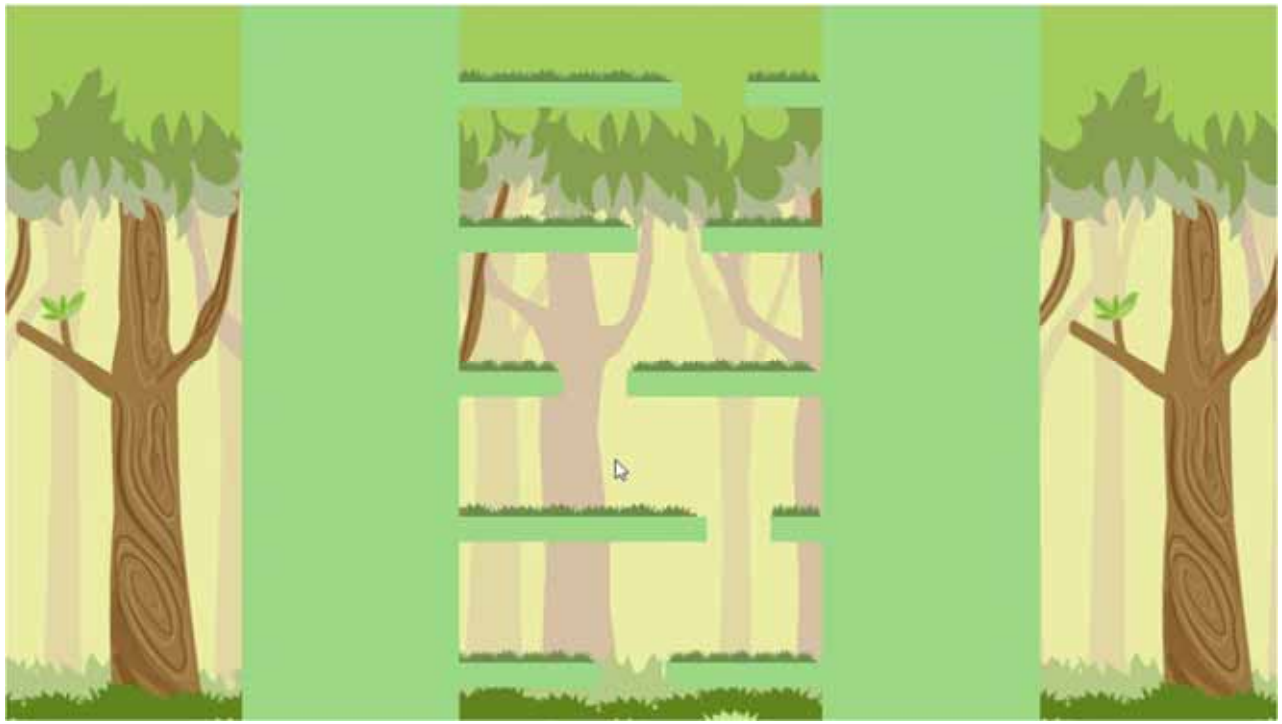
- p_{ij} … **probability** of ants that they, at a particular node i, select the route from node $i \rightarrow j$ **("heuristic desirability")**
- $\alpha > 0$ and $\beta > 0$ … the **influence parameters** (α … history coefficient, β …heuristic coefficient) usually $\alpha \approx \beta \approx 2 < 5$
- τ_{ij} … the **pheromone value** for the components, i.e. the amount of pheromone on edge (i, j)
- k … the set of usable components
- J_i … the set of nodes that ant k can reach from v_i (tabu list)
- $\eta_{ij} = \frac{1}{dij}$ … attractiveness computed by a heuristic, indicating the "a-priori **desirability"** of the move

Experiment 1: The travelling Snakesman

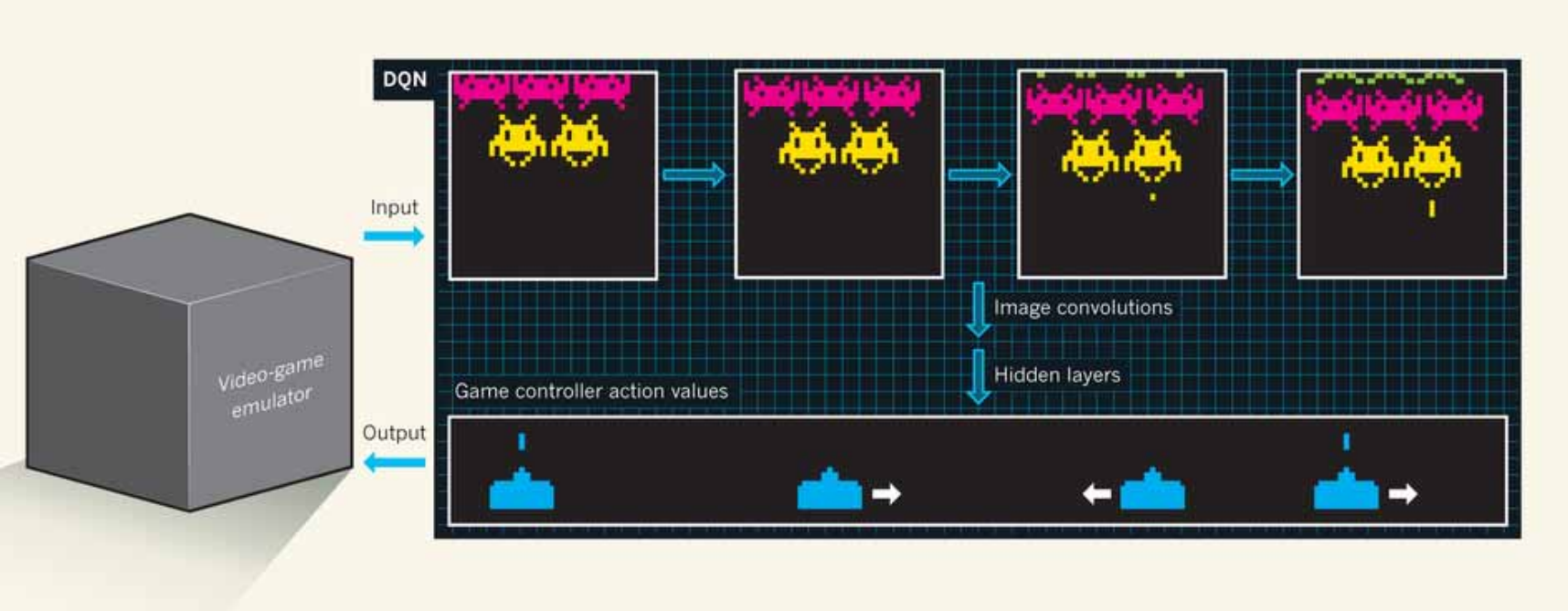
https://human-centered.ai/gamification-interactive-machine-learning/



Experiment 2: The travelling tree branch



If Google is doing their experiments with Games …



Mnih, V., Kavukcuoglu, K., Silver, D., Rusu, A. A., Veness, J., Bellemare, M. G., Graves, A., Riedmiller, M., Fidjeland, A. K., Ostrovski, G., Petersen, S., Beattie, C., Sadik, A., Antonoglou, I., King, H., Kumaran, D., Wierstra, D., Legg, S. & Hassabis, D. 2015. Human-level control through deep reinforcement learning. **Nature,** 518, (7540), 529-533, doi:10.1038/nature14236

If Google is doing his grand experiments via games, so can we … thank you!