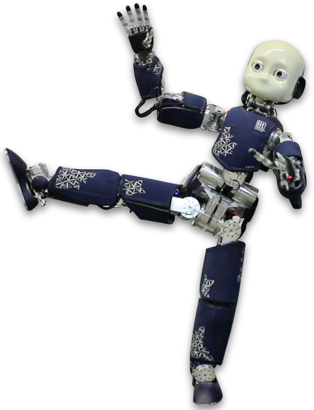
Statistics & Validation

Applied Machine & Deep Learning (190.015)

Univ.-Prof. Dr. Elmar Rueckert

Telefon: +43 3842 402 - 1901
Email: teaching@ai-lab.science



WO AUS FORSCHUNG ZUKUNFT WIRD

Chair of Cyber-Physical-Systems

Topics of this lecture

	MON 02.10.2023	TUE 03.10.2023	WED 04.10.2023	THUR 05.10.2023	FRI 06.10.2023
Topic	Intro to ML Organisation	Neural Networks	Representation Learning	Robot Learning	AML Projects
9 am					
:15					
:30					
:45					
10 am					
:15	Quizz on ML	Quizz on Neural Nets	Introduction to Deep Representation Learning		Quizz on AML
:30	Introduction to ML	Introduction to Multi-Layer-Perceptrons			Project Topic Presentations
:45					
11 am	15 min Break	15 min Break	JupyterHub NB on Rep. Learning		
:15	Statistics, Model Validation, Figures & Evaluations	Handout on Neural Networks using playground.tensorflow			Team Ass., Git Repos & Wiki Instructions
:30			30 min Break		
:45					AML Summary
12 pm	30 min Lunch Break	30 min Lunch Break	Curiosity (MLPs), Imagination (Dreamer) and Information (Empowerment)		
:15				Quizz on Robotics	
:30	Course Organisation & Grading	Introduction to CNNs		Introduction to Robot Learning	
:45					
1 pm	15 min Break	15 min Break	Quizz Summary		
:15	Python Programming with our JupyterHub	JupyterHub NB on MLPs CNNs		15 min Break	
:30				Handout on Robot Learning (Model Learning & RL)	
:45	Quizz Summary	Quizz Summary			
2 pm					
:15				15 min Break	
:30				Introduction to Mobile Robotics & SLAM	
:45					
3 pm				JupyterHub NB on Path Planning	
:15					
:30				Quizz Summary	
:45					

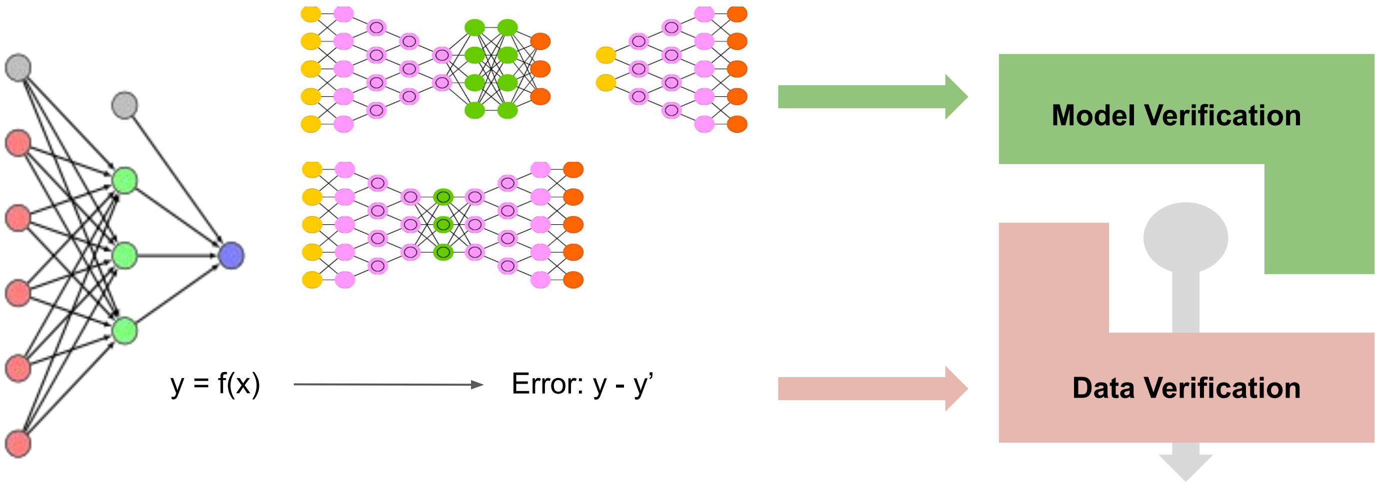
Legend

Quizz on ML	Online Quizz using https://tweedback.de
Course Content Presentation	Using google slides, etc.
15 min Break	Breaks to recover or to continue programming
Organisation & Instructions	Using google slides, etc.
Practical Exercise	Using online tools, our JupyterHub, etc.
Latest Research	State-of-the-art research

Outlook of this lecture

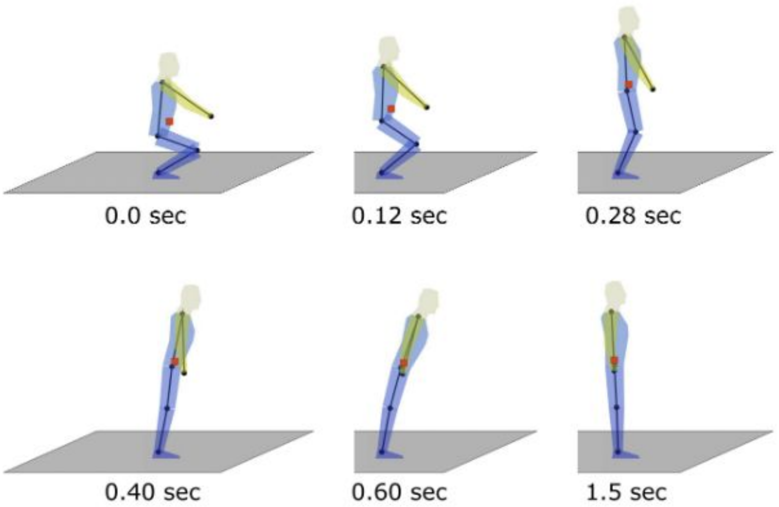
- Model Verification and Evaluation
- Overfitting & No free Lunch Theorem
- K-fold cross validation
- Figures - Best Practices & Examples

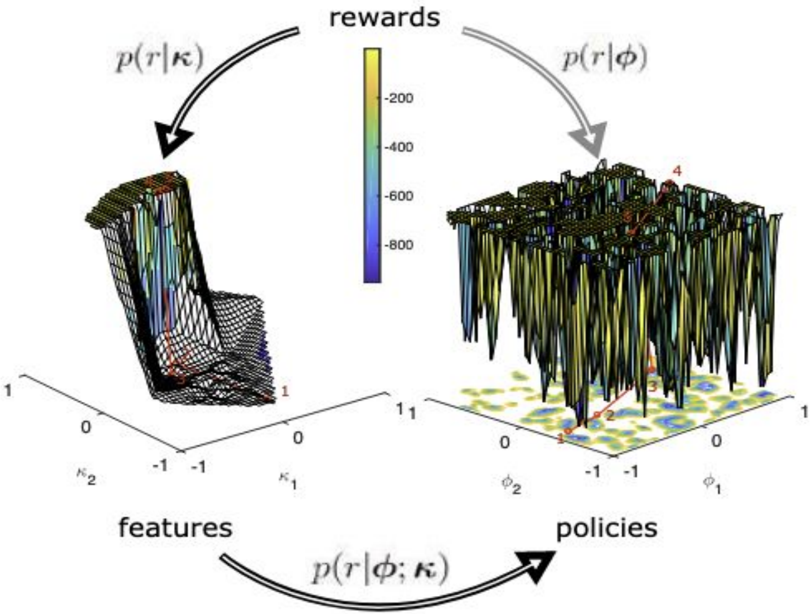
Three principled ways to verify a model



Formal Verification: Vapnik–Chervonenkis (VC) Dimension for sigmoid activation functions: at least $\Omega(|E|^2)$, at most $O(|E|^2 |V|^2)$, with the sets E of edges and V of Nodes in a directed acyclic graph.

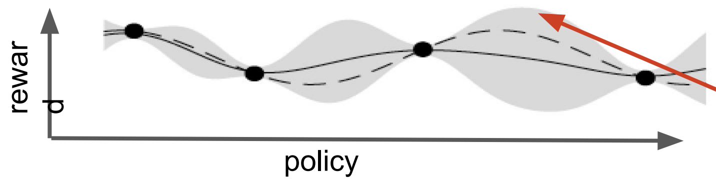
Model Verification Example - Learning to Balance





Rottmann, N.; Kunavar, T.; Babič, J.; Peters, J.; Rueckert, E. Learning Hierarchical Acquisition Functions for Bayesian Optimization International Conference on Intelligent Robots and Systems (IROS' 2020), 2020.

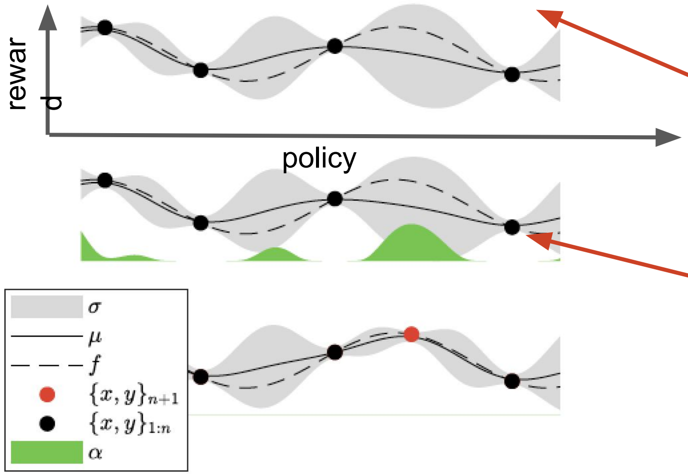
The learning algorithm is based on Bayesian Optimization (BO).
Or is a recurrent network beneficial to model the data?



BO is a global optimizer that learns a distribution over policies and rewards.

Goal: Find the best policy

The learning algorithm is based on Bayesian Optimization (BO).
Or is a recurrent network beneficial to model the data?



BO is a global optimizer that learns a distribution over policies and rewards.

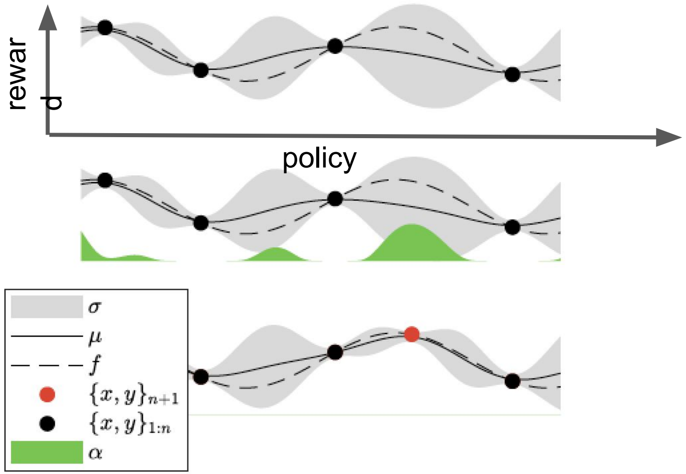
Goal: Find the best policy

BO uses an *acquisition function* to select the next best policy candidate.

e.g., the expected improvement:

$$\alpha(x; D) = (\mu(x) - \tau)\, \Phi\left(\frac{\mu(x) - \tau + \xi}{\sigma(x)}\right) + \sigma(x)\phi\left(\frac{\mu(x) - \tau + \xi}{\sigma(x)}\right),$$

The learning algorithm is based on Bayesian Optimization (BO).
Or is a recurrent network beneficial to model the data?



Our approach: Using a hierarchical acquisition function.

- Hierarchical BO: features -> policies

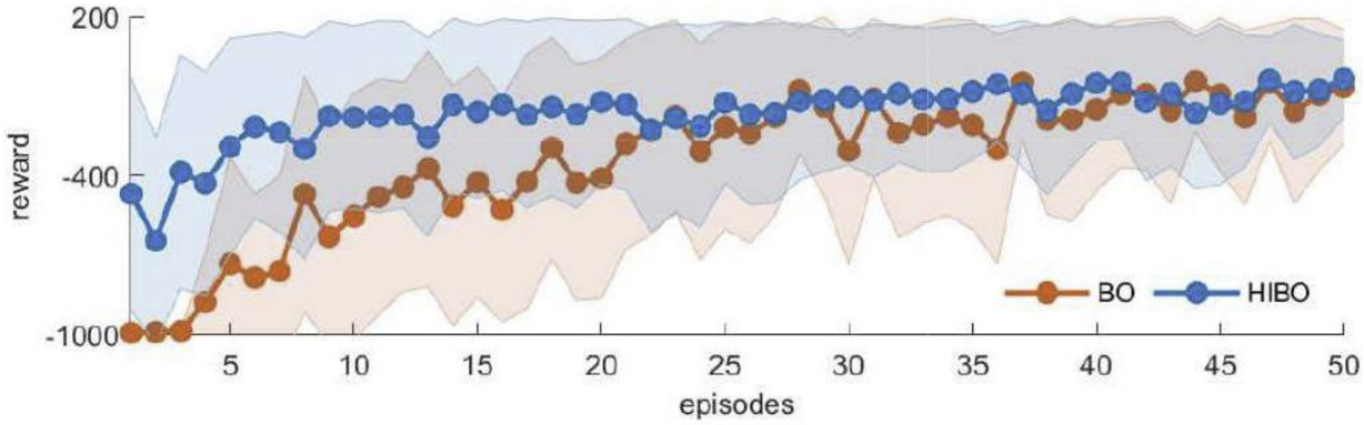
$$\boldsymbol{c}^{[k+1]} = \max_{\boldsymbol{c}} \alpha(\boldsymbol{c}; D^{[1:k]})$$

$$\boldsymbol{\theta}^{[k+1]} = \max_{\boldsymbol{\theta}} \alpha(\boldsymbol{\theta}; \boldsymbol{c}^{[k+1]}, D^{[1:k]})$$

Learning Performance using features 4 & 5 only

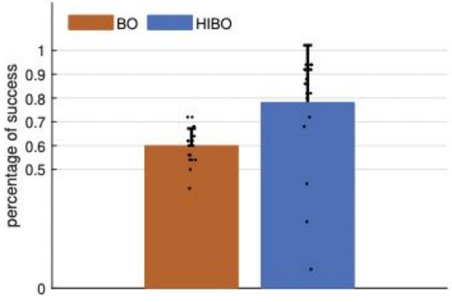
TABLE III: Feature description

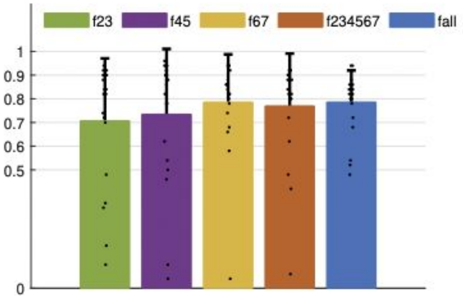
Feature 1	success
Feature 2	maximum deviation of the CoM in x direction
Feature 3	maximum deviation of the CoM in y direction
Feature 4	velocity of the CoM in x direction at 200 ms
Feature 5	velocity of the CoM in y direction at 200 ms
Feature 6	velocity of the CoM in x direction at 400 ms
Feature 7	velocity of the CoM in y direction at 400 ms

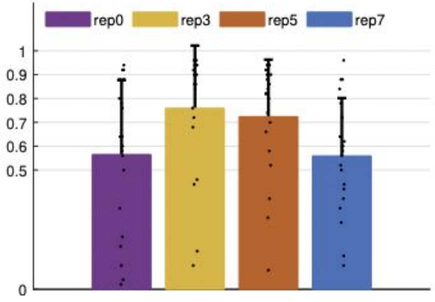


Model Verification Example - Learning to Balance

Learning Performance using different selections of features – the model verification!



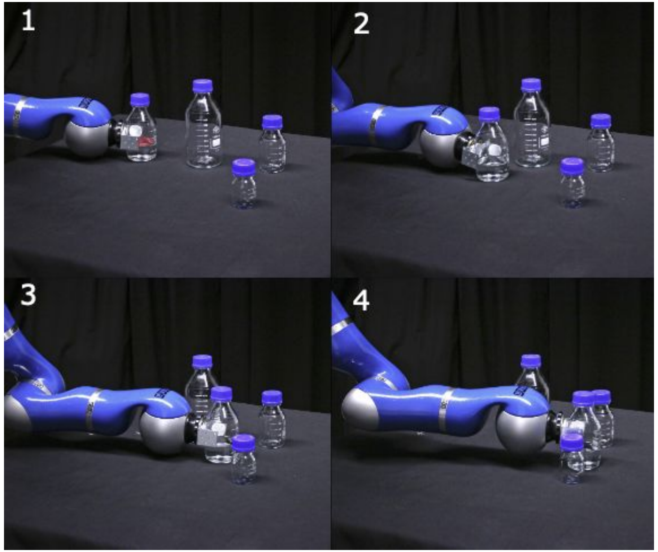


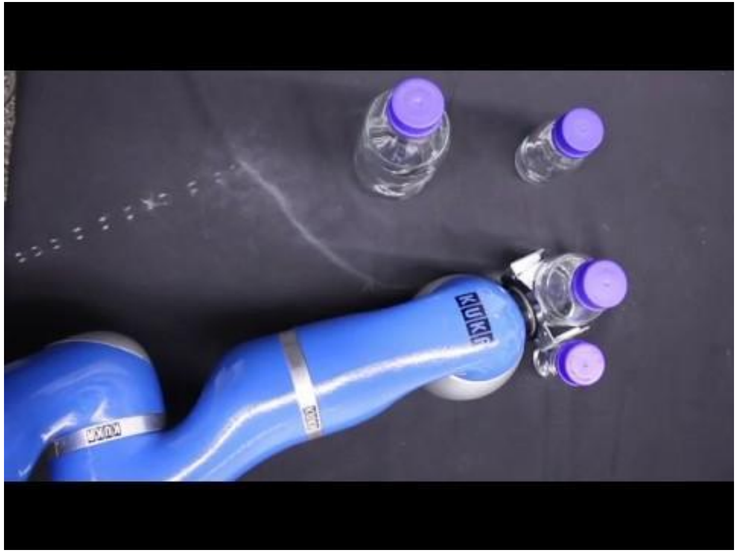


Note: Mean and standard deviations are illustrated **and** the **performance scores of the individual experiments** as dots. The scattering of the dots indicates how often the RL approach gets stuck in a **sub-optimal local minima**.

Rottmann, N.; Kunavar, T.; Babič, J.; Peters, J.; Rueckert, E. Learning Hierarchical Acquisition Functions for Bayesian Optimization International Conference on Intelligent Robots and Systems (IROS' 2020), 2020.

Data Verification Example - Learning the Dynamics Model of a Robot arm





Rueckert, Elmar; Nakatenus, Moritz; Tosatto, Samuele; Peters, Jan [Learning Inverse Dynamics Models in O(n) time with LSTM networks](#)
Proceedings of the International Conference on Humanoid Robots (HUMANOIDS), 2017.

Data Verification Example - Learning the Dynamics Model of a Robot arm

The goal is to predict the optimal motor torques given the joint angles, velocities and accelerations.

This function is called **Inverse Dynamics Model:**

$$\boldsymbol{\tau} = M(\boldsymbol{q})\ddot{\boldsymbol{q}} + h(\boldsymbol{q}, \dot{\boldsymbol{q}}) + \epsilon(\boldsymbol{q}, \dot{\boldsymbol{q}}, \ddot{\boldsymbol{q}})$$

A supervised learning problem:

$$\boldsymbol{y} = f(\boldsymbol{x}) + \boldsymbol{\zeta} : \mathbb{R}^{3d \times 1} \mapsto \mathbb{R}^{d \times 1}$$

$$\boldsymbol{y} = \boldsymbol{\tau}$$

$$\boldsymbol{x} = [\boldsymbol{q}^T, \dot{\boldsymbol{q}}^T, \ddot{\boldsymbol{q}}^T]^T$$

Given a **dataset**, the goal is to minimize the **mean-squared error** (**MSE)** of the predicted and the true motor torques:

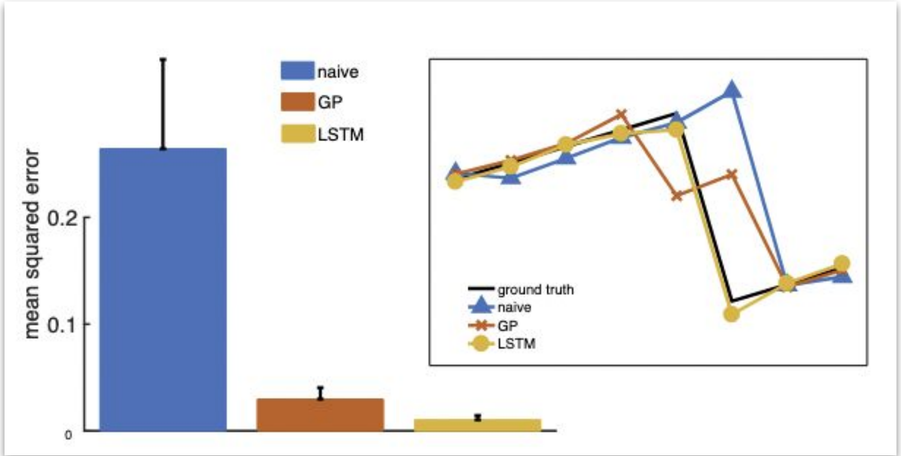
$$\text{dataset } \mathfrak{D} = \langle \boldsymbol{x_t}, \boldsymbol{y_t} \rangle_{t=1,\ldots,n}$$

$$MSE = \frac{1}{d\,n} \sum_{j=1}^{d} \sum_{t=1}^{n} \left(\hat{y}_t^{[j]} - \tilde{y}_t^{[j]} \right)^2$$

Data Verification Example - Learning the Dynamics Model of a Robot arm

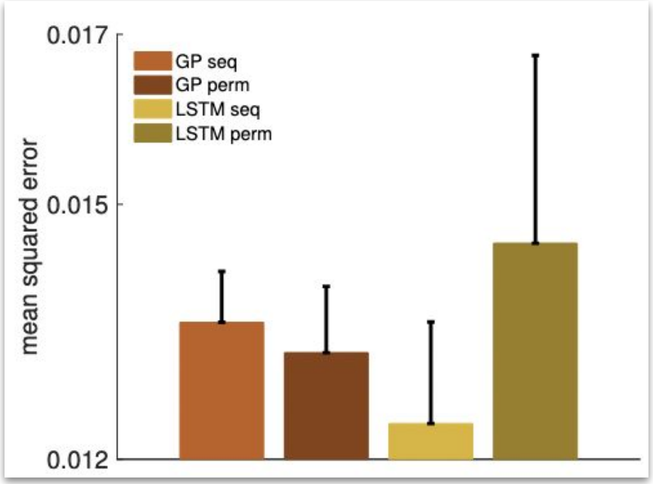
Verification of the Learning Algorithm - Baseline Comparisons

- The naive approach is to simply compute the output based on the sum of the current state and the current velocity.
- Gaussian Processes are state-of-the-art model learning approaches.
- Our approach is a recurrent neural network called long-short-term-memory (LSTM) network.



Data Verification Example - Learning the Dynamics Model of a Robot arm

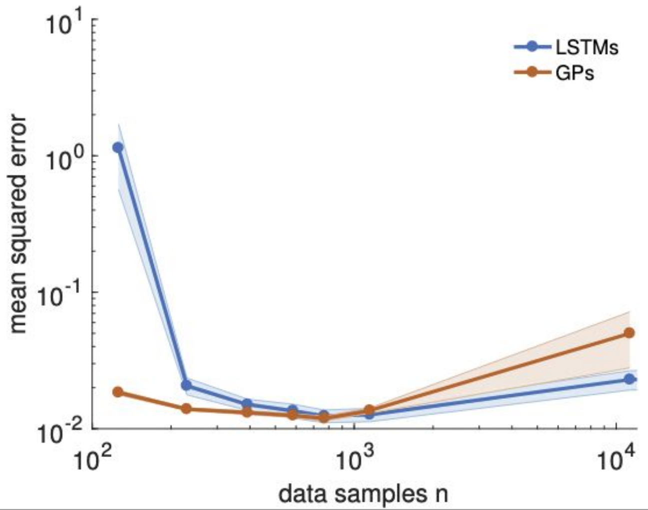
Verification of the Importance of Capturing Temporal Dependencies in the Data.
Or is a recurrent network beneficial to model the data?



Data Verification Example - Learning the Dynamics Model of a Robot arm

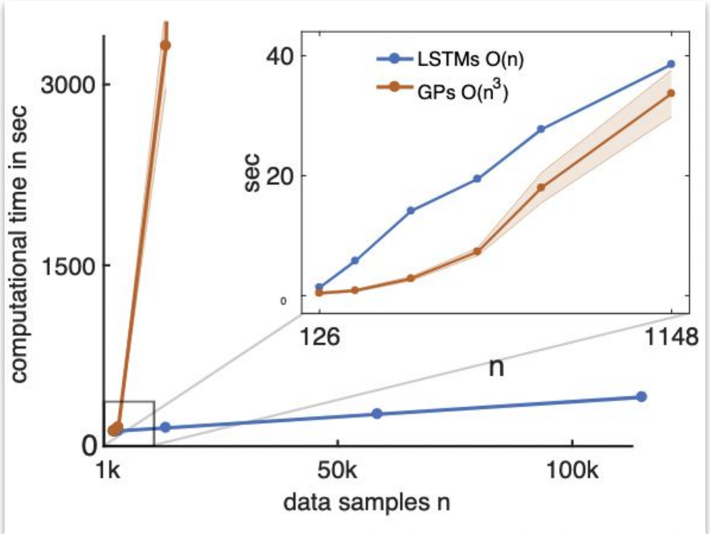
Verification of the Number of Data Samples needed for training the model.

	1k	10k	50k	100k
GPs	13.6 ± 0.6	49.9 ± 22.0	—	—
LSTMs	12.6 ± 1.4	22.9 ± 3.7	24.5 ± 4.4	21.8 ± 3.9



Data Verification Example - Learning the Dynamics Model of a Robot arm

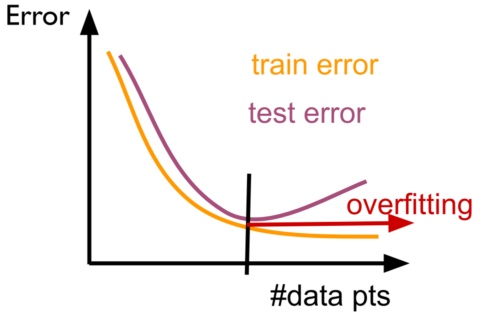
Evaluation of the Computational & Memory Costs of the model.

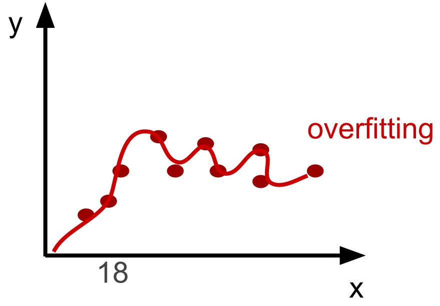


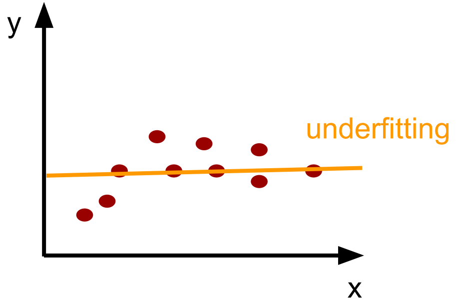
Rueckert, Elmar; Nakatenus, Moritz; Tosatto, Samuele; Peters, Jan [Learning Inverse Dynamics Models in O(n) time with LSTM networks]
Proceedings of the International Conference on Humanoid Robots (HUMANOIDS), 2017.

Overfitting & No Free Lunch Theorem

By that we can detect Overfitting or Underfitting!





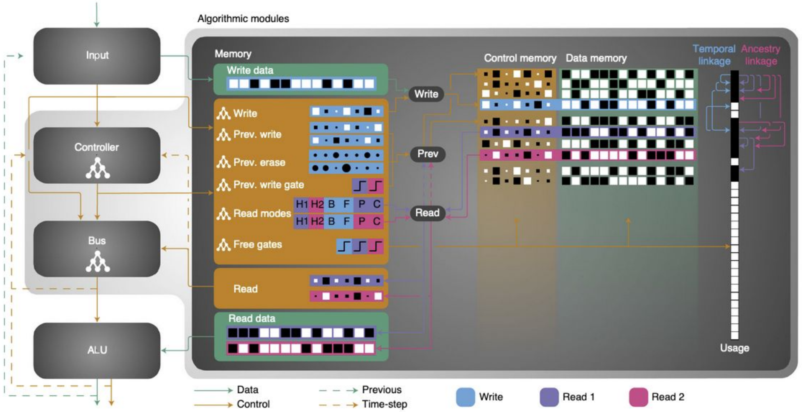


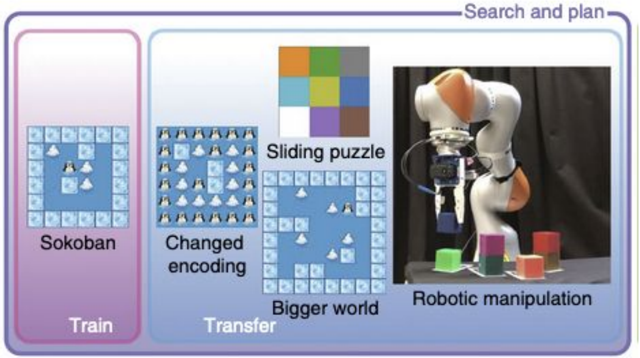
Avoid underfitting: As long as you do not observe an increase in the test error, you should increase the model complexity! For example, increase the number of layers or neurons in a neural network.

No free Lunch Theorem

“All models are wrong, but some models are useful”, by George Box (Box and Draper 1987, p424).

There is not a single best model or learning algorithm that works optimally for all kinds of problems. This is called the no free lunch theorem (Daniel Wolpert 1996).

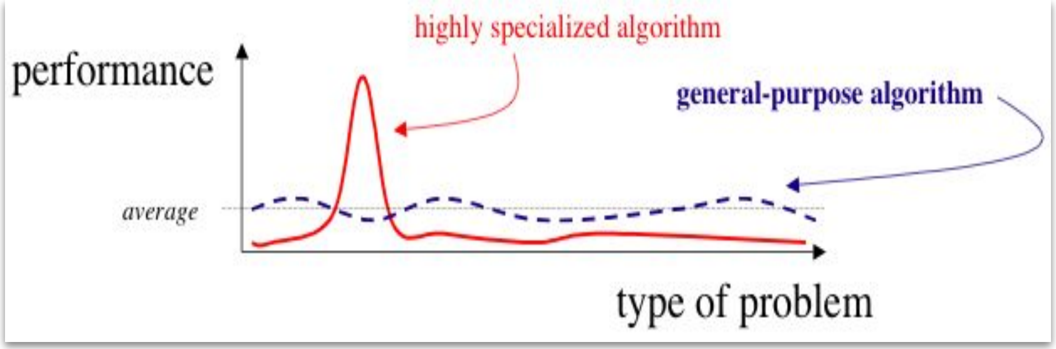




Example of a Cross-domain research study: Tanneberg, Daniel; Rueckert, Elmar; Peters, Jan. Evolutionary training and abstraction yields algorithmic generalization of neural computers, Nature Machine Intelligence, pp. 1–11, 2020.

No free Lunch Theorem

"As a consequence of the no free lunch theorem, we need to develop many different types of models, to cover a wide variety of data that occurs in the real world", by Kevin P. Murphy (p24 in his machine learning book)



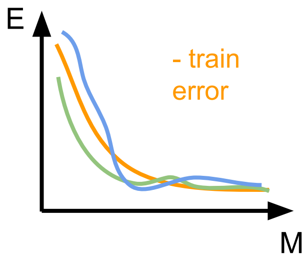
From https://medium.com/@LeonFedden/the-no-free-lunch-theorem-62ae2c3ed10c

K-fold cross validation

- To verify the correctness of the model on new but related data.
- To test the generalization ability.
- To identify Overfitting or Underfitting.

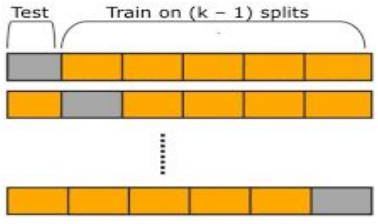
- **Ideal case: Large Dataset**

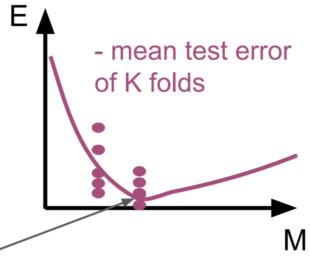
Dataset Splitt



- **Typical case: Small Datasets** (or long tails)

K-Fold Cross Validation

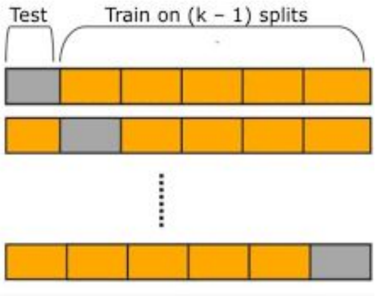


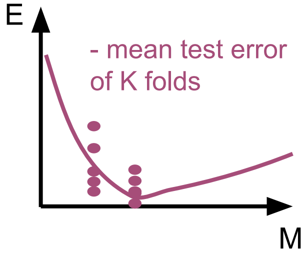


Each dot corresponds to one test result. Compute the mean and the standard deviation to identify when Overfitting kicks in.

K-fold Cross Validation

- **Data Verification:** To test if your dataset is large enough.
- **Model Verification:** To test if you model has enough free parameters.





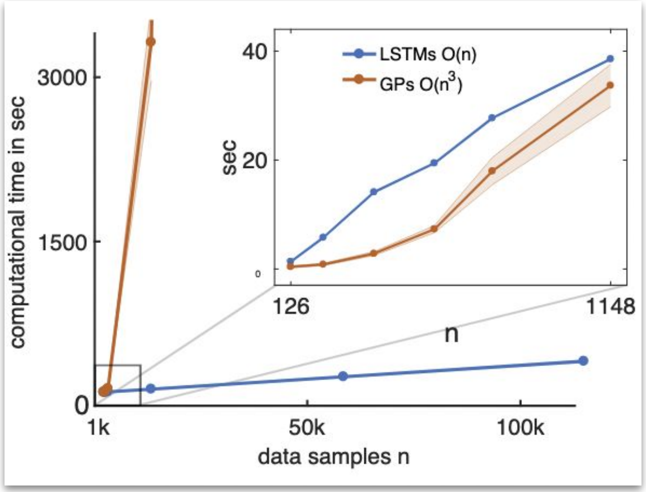
Examples:

- **Data Verification**:
  - M … samples: Test for the minimum number of training samples,
  - M … samples: Are the data sets balanced?

- **Model Verification**:
  - M … network or model type.
  - Number of Hidden Layer, neurons etc.

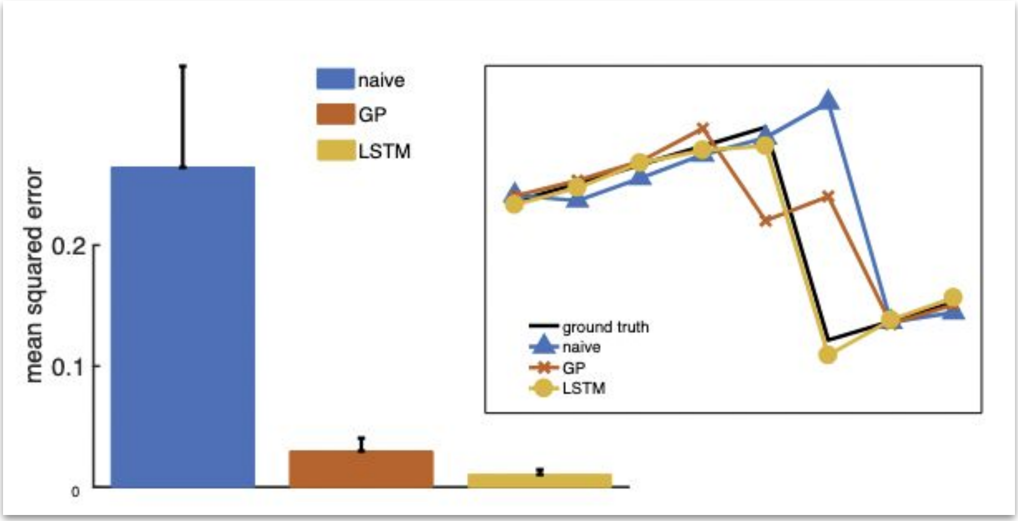
Figures

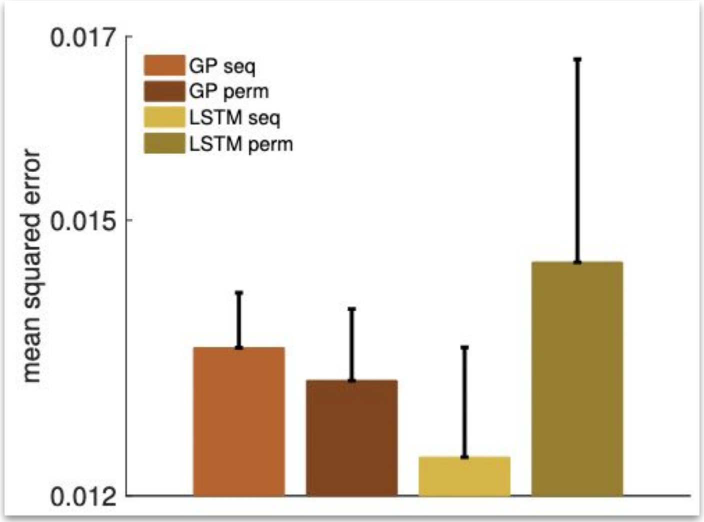
- Make sure that all axis have labels.
- Always add a legend (images or tables may be exceptions).
- Use different line colors, e.g., via matplotlib.colors import LinearSegmentedColormap, ListedColormap.
- Use a minimum line width of 2 and different line styles.
- The font size of text in figures should be equal to the figure caption font size.
- The caption of a figure needs to be self-explaining. All major elements in a figure need to be defined.
- Remove bounding boxes of legends and of the graph if not needed.
- Remove all redundant or unnecessary elements.



Bar Plot Examples

With an inline detail plot.

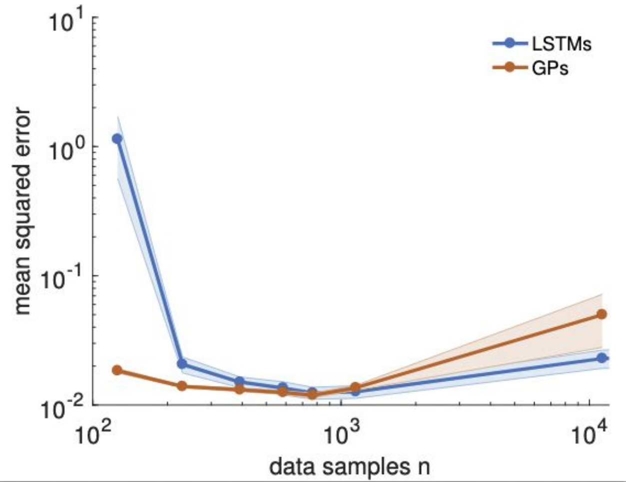


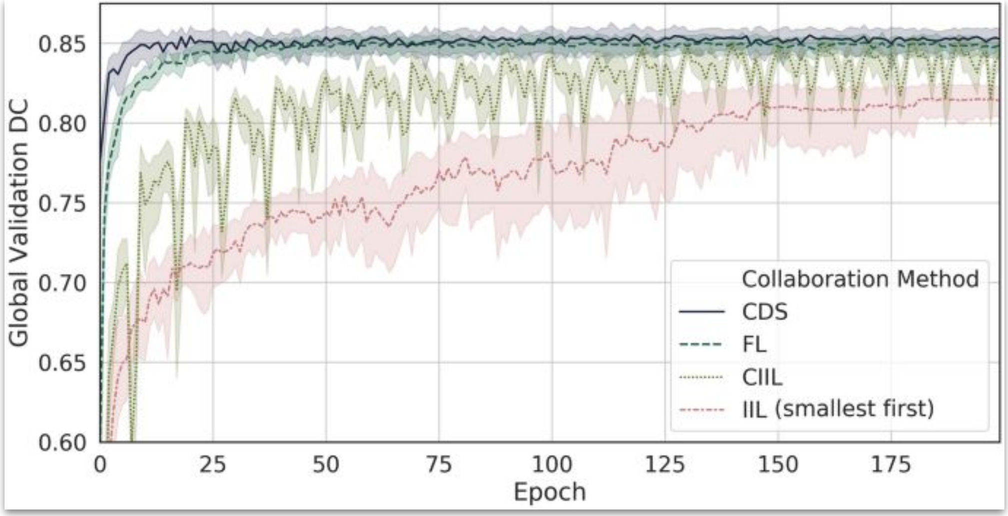


Learning Curve Examples

With a statistics as a table on top.

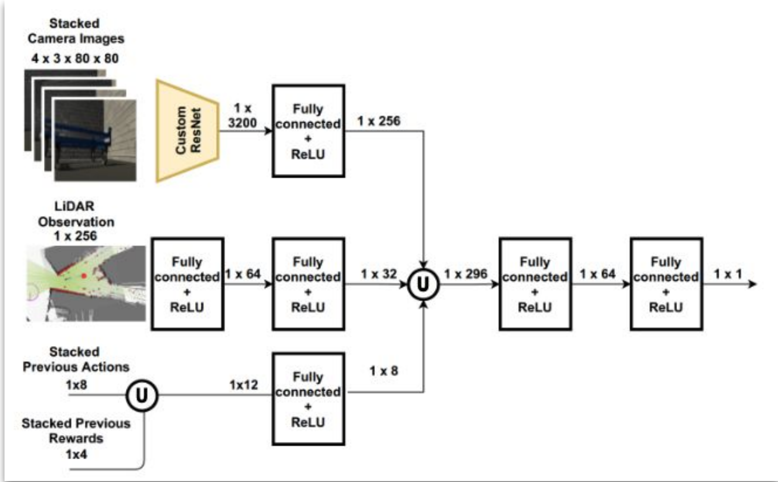
	1k	10k	50k	100k
GPs	13.6 ± 0.6	49.9 ± 22.0	—	—
LSTMs	12.6 ± 1.4	22.9 ± 3.7	24.5 ± 4.4	21.8 ± 3.9

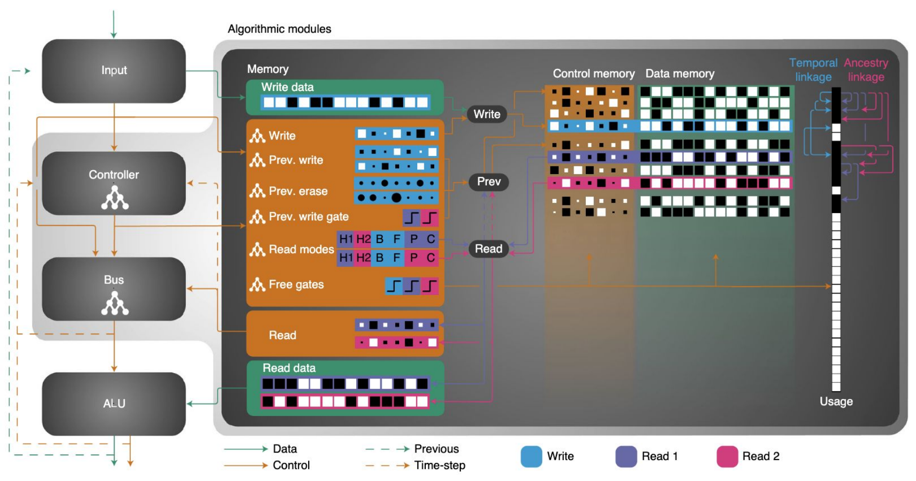




Typical evaluation of multiple learning methods on multiple ‘runs’, from [Sheller et al. 2020](#).

Neural Network Model Figures

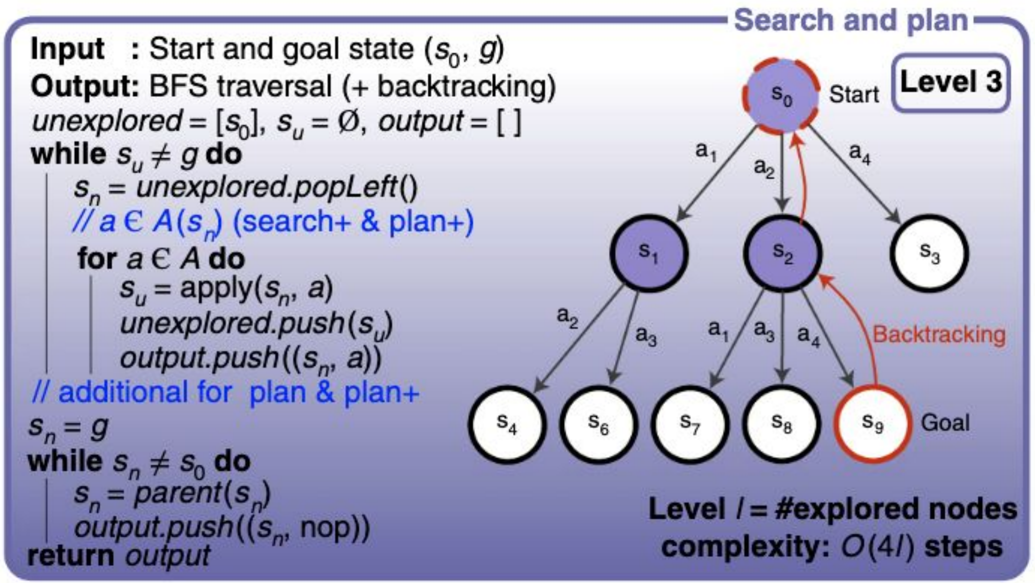




Xue, Honghu; Hein, Benedikt; Bakr, Mohamed; Schildbach, Georg; Abel, Bengt; Rueckert, Elmar. **Using Deep Reinforcement Learning with Automatic Curriculum Learning for Mapless Navigation in Intralogistics**. Applied Sciences (MDPI), Special Issue on Intelligent Robotics, 2022.

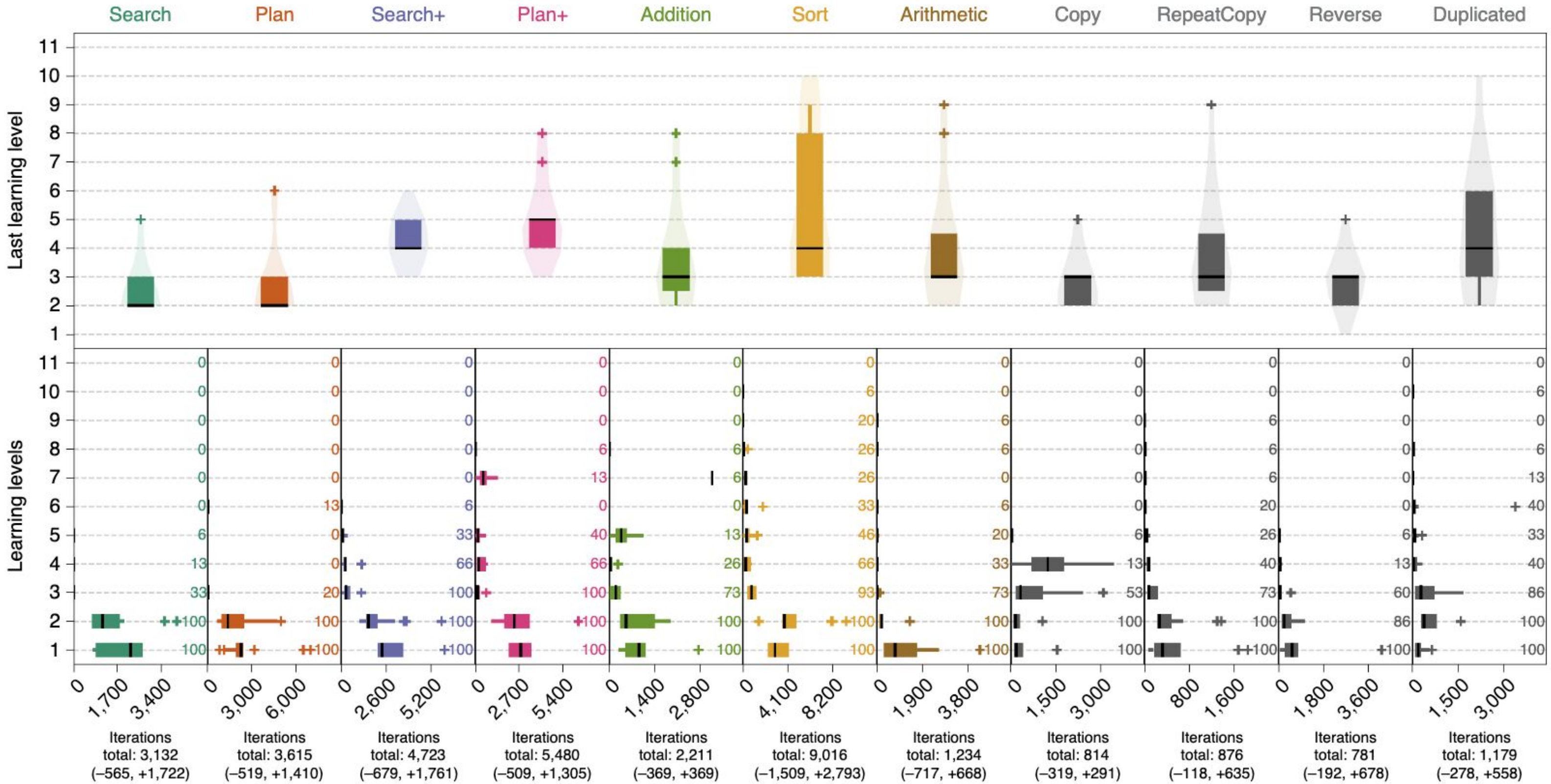
Tanneberg, Daniel; Rueckert, Elmar; Peters, Jan. **Evolutionary training and abstraction yields algorithmic generalization of neural computers.** Nature Machine Intelligence, pp. 1–11, 2020.

Algorithm & Task Figures



Tanneberg, Daniel; Rueckert, Elmar; Peters, Jan. **Evolutionary training and abstraction yields algorithmic generalization of neural computers**. Nature Machine Intelligence, pp. 1–11, 2020.

Task Performance Statistics



Tanneberg, Daniel; Rueckert, Elmar; Peters, Jan. **Evolutionary training and abstraction yields algorithmic generalization of neural computers.** Nature Machine Intelligence, pp. 1–11, 2020.

More on Figures will follow in the Jupyter Python Tutorial

```
%%html •••
```

CPS

Chair of Cyber-Physical-Systems, Austria

Credentials	
Host	Montanuniversitaet Leoben
Web	https://cps.unileoben.ac.at
Mail	cps@unileoben.ac.at
Authors	Vedant Dave, Fotis Lygerakis, Linus Nwankwo, Melanie Neubauer, Nikolaus Feith and Elmar Rueckert
Corresponding Authors	melanie.neubauer@unileoben.ac.at, rueckert@unileoben.ac.at
Last edited	02.10.2023

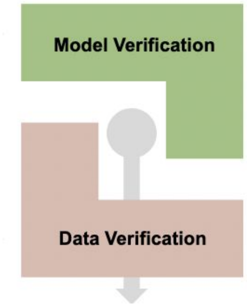
Database Basics

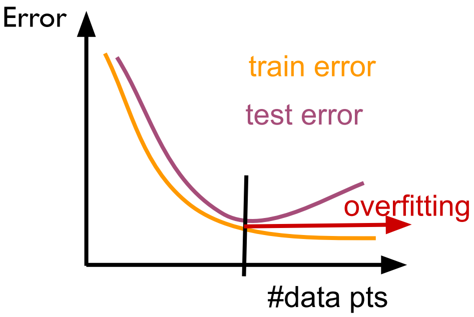
This is a tutorial on the basics of interfacing databases for beginners.

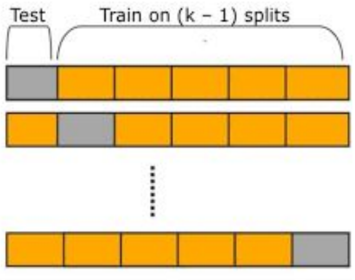
Content

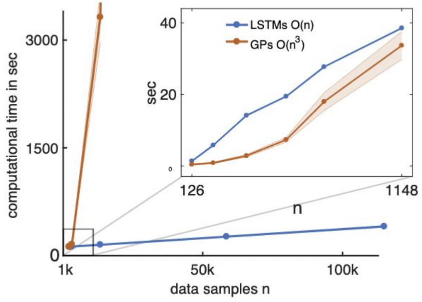
Summary of Statistics & Validation

- Model Verification and Evaluation
- Overfitting & No free Lunch Theorem
- K-fold cross validation
- Figures - Best Practices & Examples



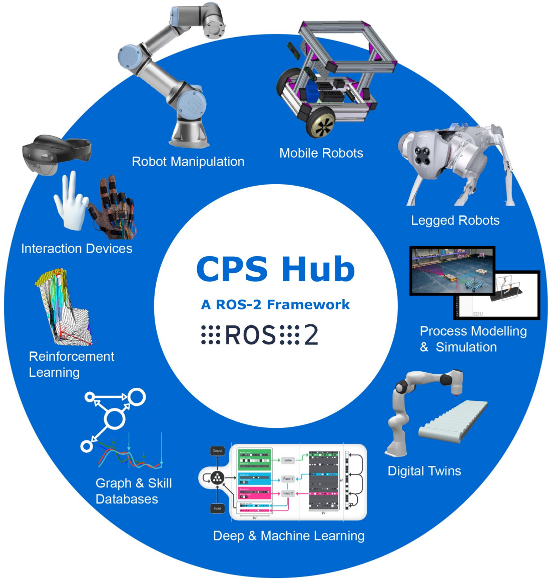






Thank you for your attention!

Visit our Youtube Channel:
https://youtube.com/@CPSAustria



Phone: +43 3842 402 – **1901** (Sekretariat CPS)

Email: cps@unileoben.ac.at

Web: https://cps.unileoben.ac.at