Trends and Challenges in Machine Learning

Iuri Rocha, TUD

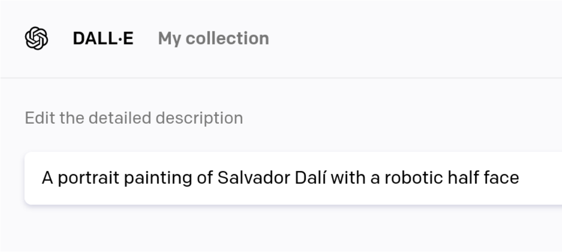
Michael Abdelmalik, TU/e

Hongyang Cheng, UT

Francesco Maresca, RUG

But first...

Trends and challenges in mechanics, according to machine learning:



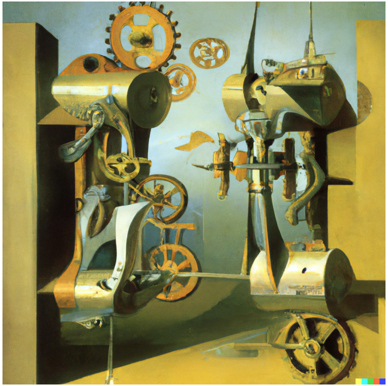
A portrait painting of Salvador Dalí with a robotic half face

But first...

Trends and challenges in mechanics, according to machine learning:

A complicated Finite Element model, by Vincent van Gogh

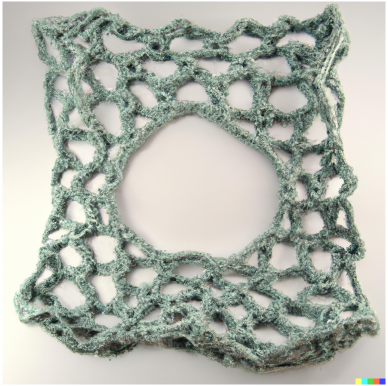
Engineering Mechanics, by Salvador Dalí



But first...

Trends and challenges in mechanics, according to machine learning:

A Finite Element model knitted out of wool



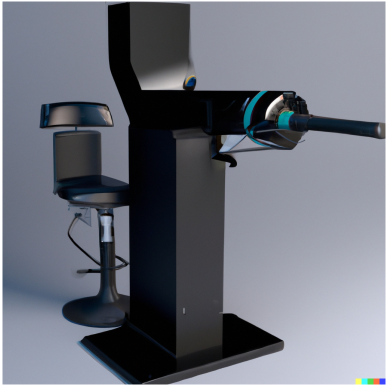
Two teddy bears discovering a new metamaterial

But first...

Trends and challenges in mechanics, according to machine learning:

An ancient Greek running a computer model, marble, ca. 200 BC

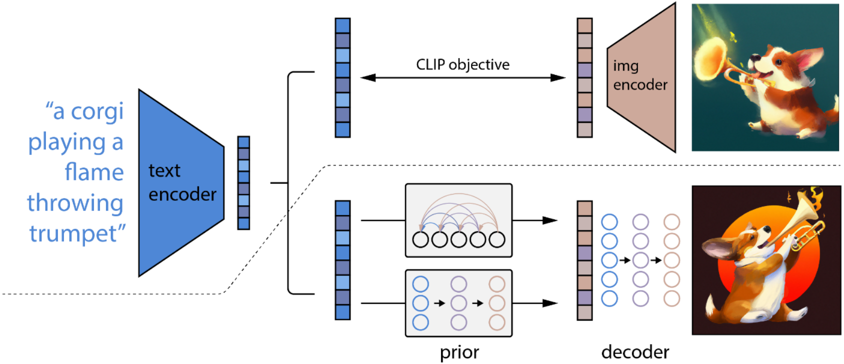
A futuristic Universal Testing Machine



Taking a peek under the hood

Combination of encoder-decoder architectures, multistage training:

- 3.5 billion parameters, trained with about 600 million text-image pairs
- A Gaussian diffusion decoder constructs new images every forward pass, pixel by pixel

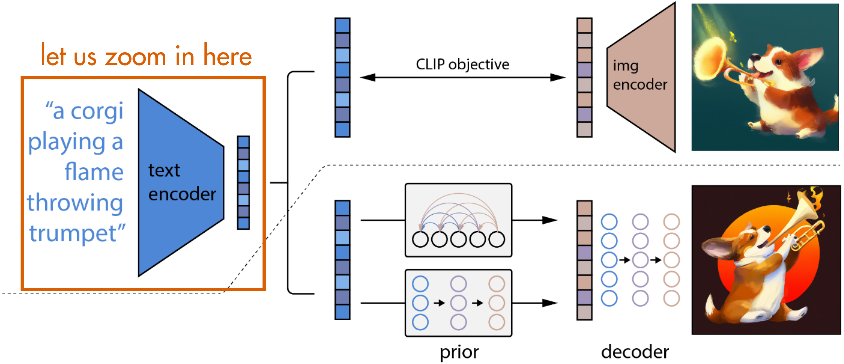


Taking a peek under the hood

Combination of encoder-decoder architectures, multistage training:

- 3.5 billion parameters, trained with about 600 million text-image pairs
- A Gaussian diffusion decoder constructs new images every forward pass, pixel by pixel

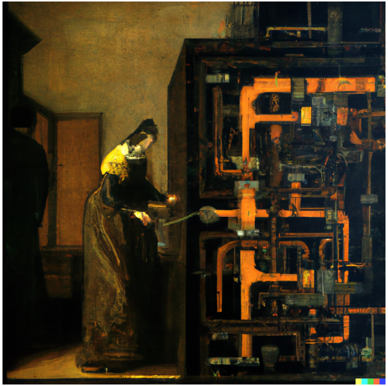
let us zoom in here

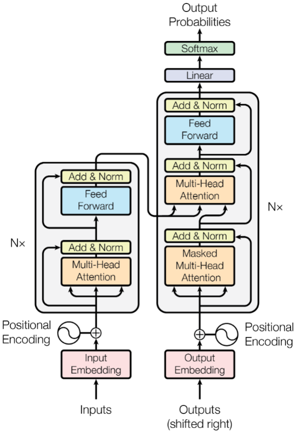


Taking a peek under the hood

A Transformer mapping word sequences to latent sequences:

A Transformer neural network, by Johannes Vermeer



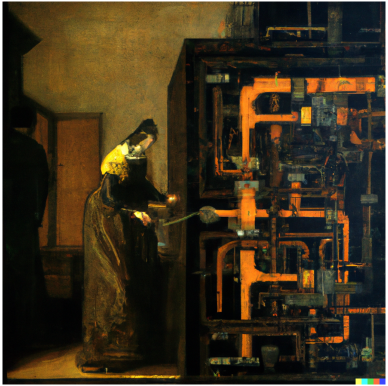


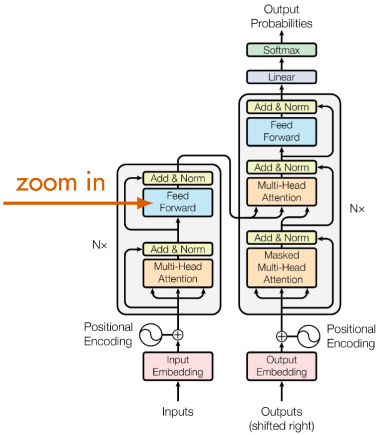
[Vaswani et al (2017), arXiv:1706.03762v5]

Taking a peek under the hood

A Transformer mapping word sequences to latent sequences:

A Transformer neural network, by Johannes Vermeer



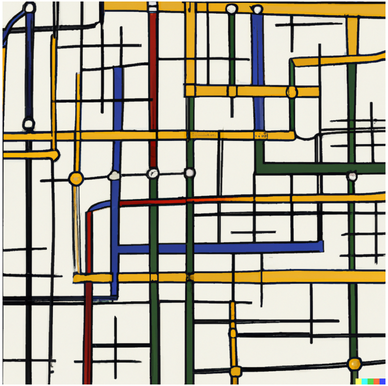


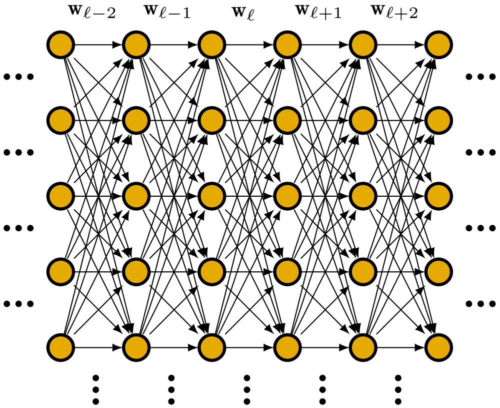
[Vaswani et al (2017), arXiv:1706.03762v5]

Taking a peek under the hood

A feedforward neural network mapping inputs to outputs:

A feedforward neural network, by Piet Mondriaan



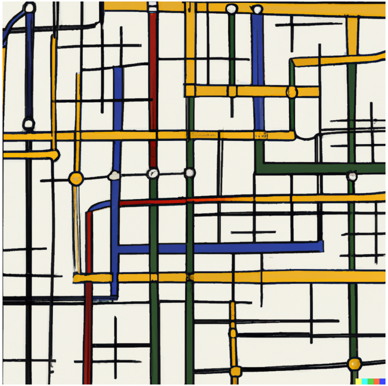


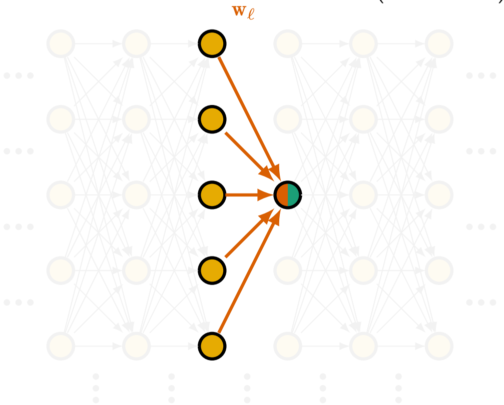
Taking a peek under the hood

A feedforward neural network mapping inputs to outputs:

$$a_{\ell+1}^{(i)} = h\left(\sum_j w_\ell^{(ij)} a_\ell^{(j)}\right)$$

A feedforward neural network, by Piet Mondriaan

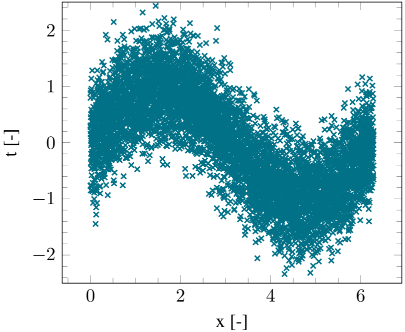




The blessing and the curse of unlimited flexibility

What to do with so many parameters? A simple example:

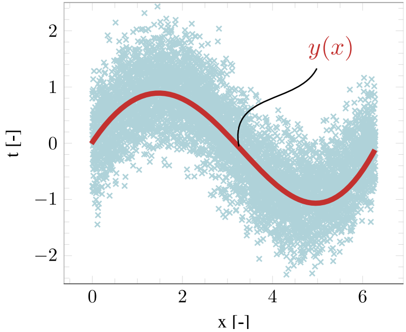
- Fitting some noisy response with a k-Nearest Neighbors model

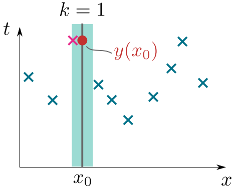


The blessing and the curse of unlimited flexibility

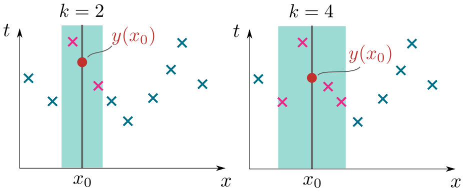
What to do with so many parameters? A simple example:

- Fitting some noisy response with a k-Nearest Neighbors model





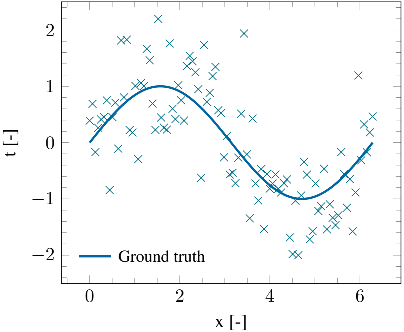
$$y(x_0) = \frac{1}{k} \sum_{x_i \in N_k} t_i$$



The blessing and the curse of unlimited flexibility

What to do with so many parameters? A simple example:

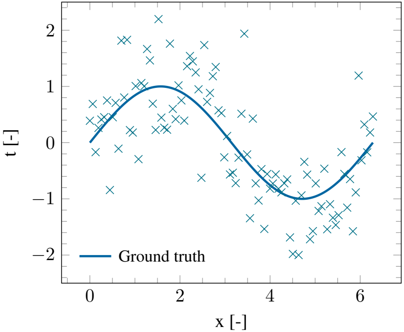
- Fitting some noisy response with a k-Nearest Neighbors model



The blessing and the curse of unlimited flexibility

What to do with so many parameters? A simple example:

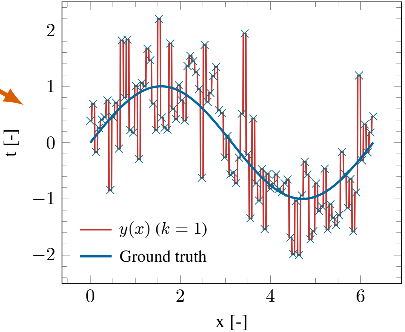
- Fitting some noisy response with a k-Nearest Neighbors model



$$y(x) = \frac{1}{k} \sum_{x_i \in N_k} t_i$$

$$L = \frac{1}{N} \sum_{i}^{N} \left(y(x_i, k) - t_i\right)^2$$

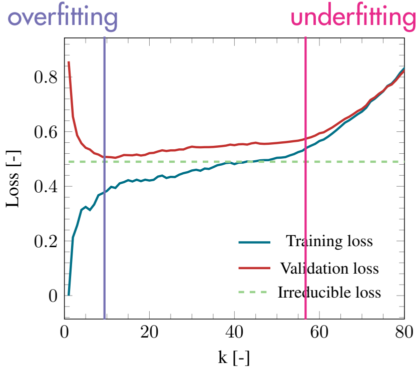
$$k = 1 \Rightarrow L = 0 \ \textbf{(!)}$$



The blessing and the curse of unlimited flexibility

The bias-variance tradeoff:

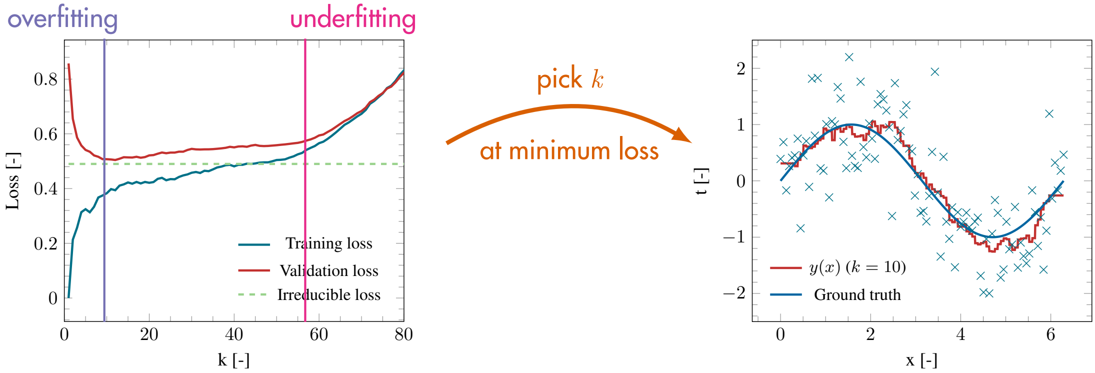
- We hide some data from the model and trade some flexibility for robustness



The blessing and the curse of unlimited flexibility

The bias-variance tradeoff:

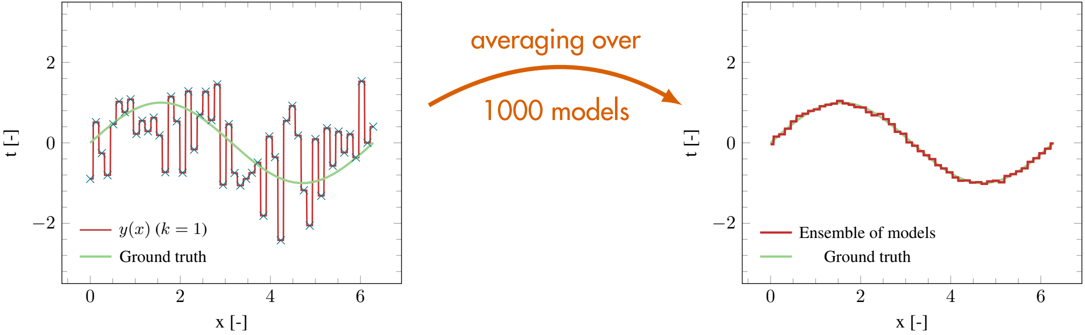
- We hide some data from the model and trade some flexibility for robustness



The blessing and the curse of unlimited flexibility

The power of averaging:

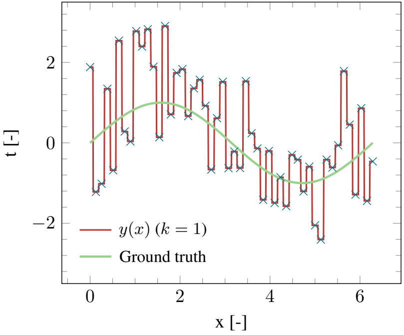
- Bayesian machine learning: uncertainty over dataset $\Rightarrow$ uncertainty over model



The blessing and the curse of unlimited flexibility

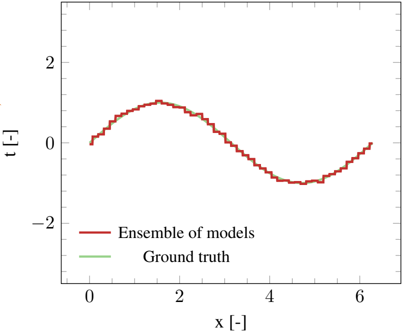
The power of averaging:

- Bayesian machine learning: uncertainty over dataset $\Rightarrow$ uncertainty over model



averaging over

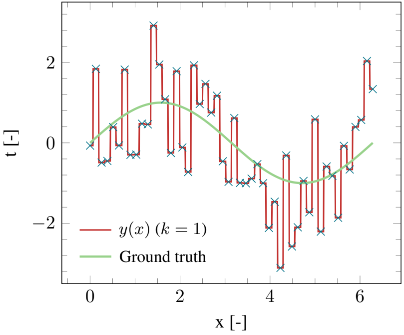
1000 models



The blessing and the curse of unlimited flexibility

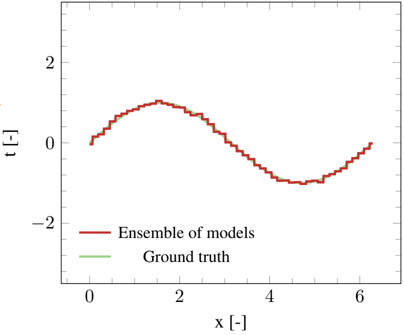
The power of averaging:

- Bayesian machine learning: uncertainty over dataset $\Rightarrow$ uncertainty over model



averaging over

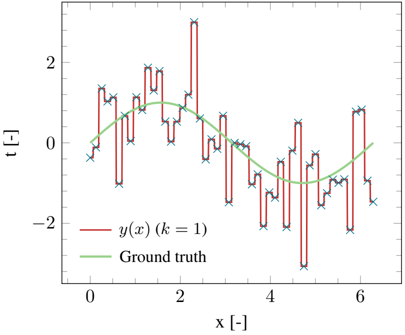
1000 models



The blessing and the curse of unlimited flexibility

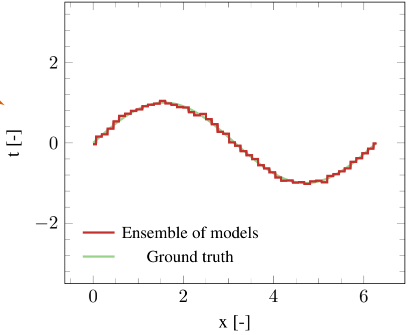
The power of averaging:

- Bayesian machine learning: uncertainty over dataset $\Rightarrow$ uncertainty over model



averaging over

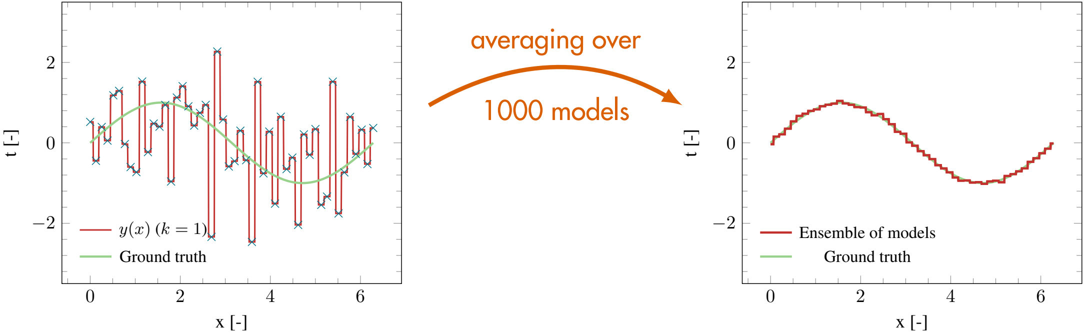
1000 models

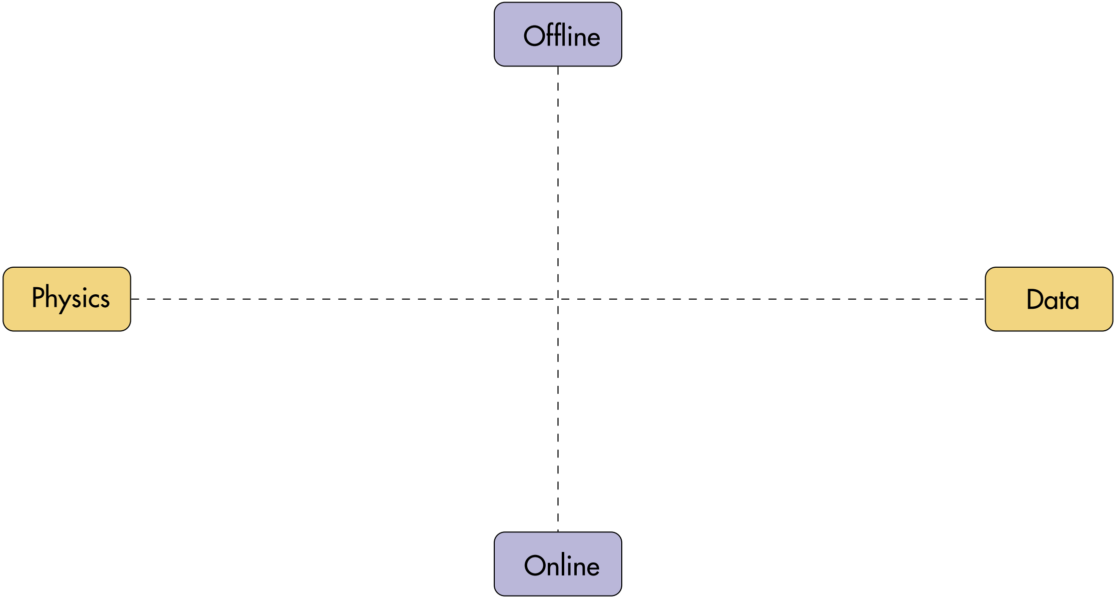


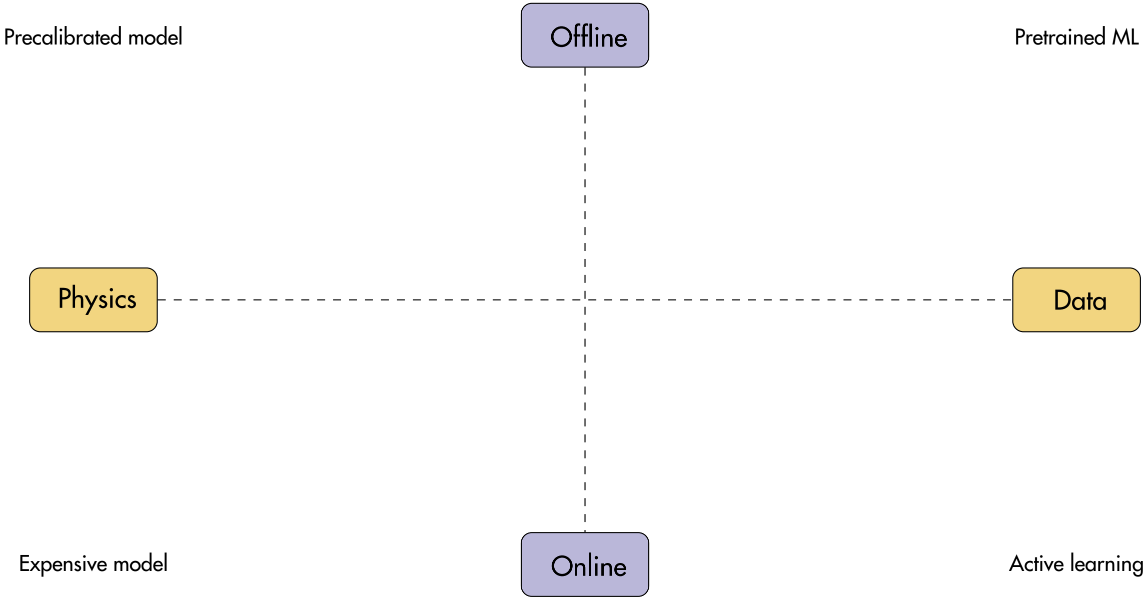
The blessing and the curse of unlimited flexibility

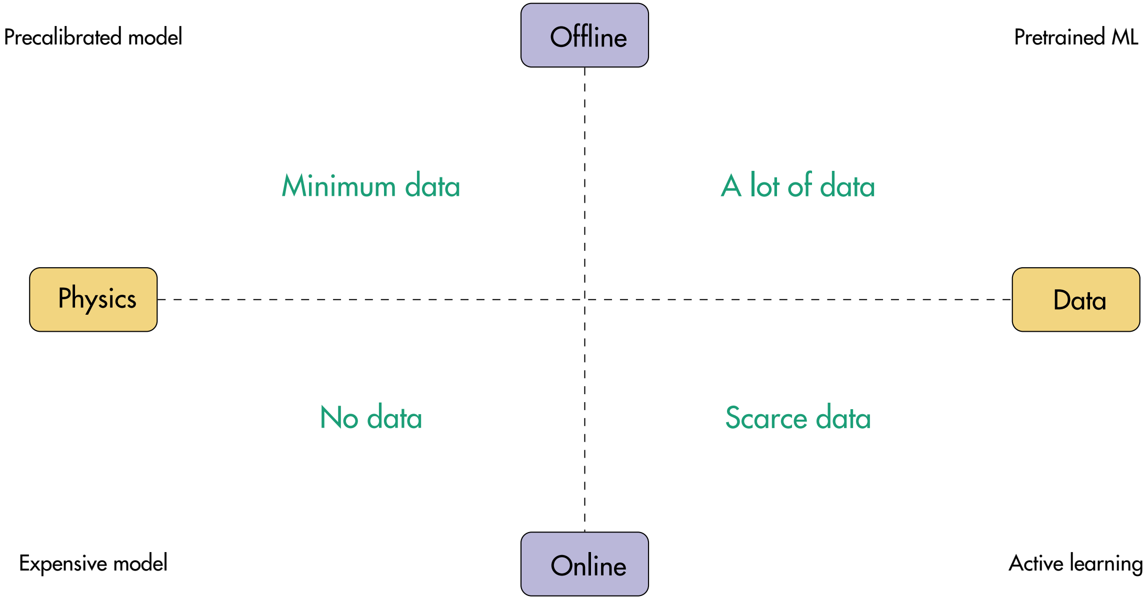
The power of averaging:

- Bayesian machine learning: uncertainty over dataset $\Rightarrow$ uncertainty over model




Offline
Physics
Data
Online


Precalibrated model
Offline
Pretrained ML
Physics
Data
Expensive model
Online
Active learning


Precalibrated model
Offline
Pretrained ML
Minimum data
A lot of data
Physics
Data
No data
Scarce data
Expensive model
Online
Active learning

Offline

Isaac Newton training a machine learning model

Physics ? →  ← ? Data

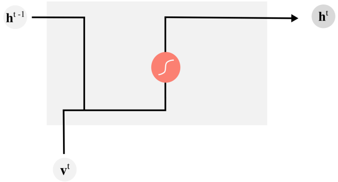
Online

Offline
Physics Data
Online

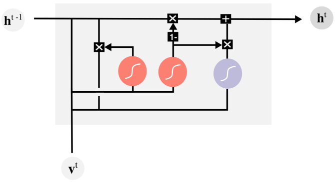
Path-dependent material behavior

Recurrent Neural Networks (RNN) for strain path dependency:

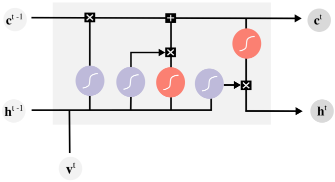
- Network includes latent (hidden) variables accounting for history dependency
- Fast surrogates for expensive models. Accurate as long as trained with enough data



Classical RNN



Gated Recurrent Unit (GRU)



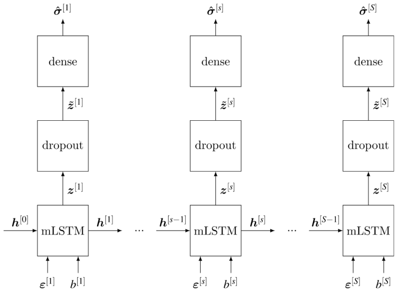
Long Short Term Memory (LSTM)

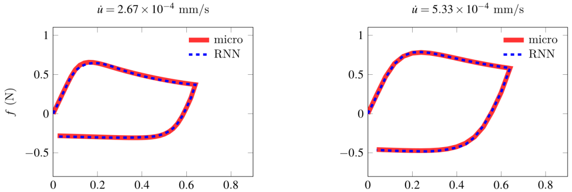
[Maia et al (2022), arXiv:2209.07320v1]

Path-dependent material behavior

Recurrent Neural Networks (RNN) for strain path dependency:

- Network includes latent (hidden) variables accounting for history dependency
- Fast surrogates for expensive models. Accurate as long as trained with enough data





[Ghavamian and Simone (2019), CMAME 357:112594]

[Angelo Simone's group (TUD)]

Offline

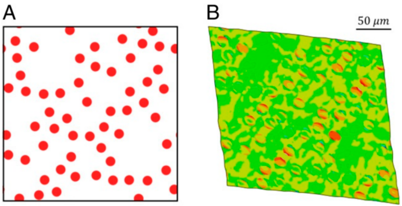
Physics Data

Online

Path-dependent material behavior

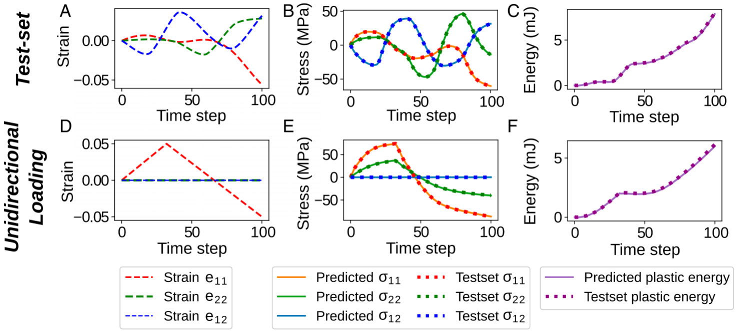
Recurrent Neural Networks (RNN) for strain path dependency:

- Network includes latent (hidden) variables accounting for history dependency
- Fast surrogates for expensive models. Accurate as long as trained with enough data



[Mozzafar et al (2019), PNAS 116:26414-26420]

[Miguel Bessa's group (TUD)]

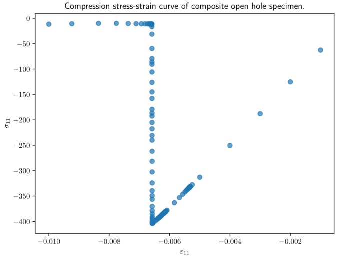


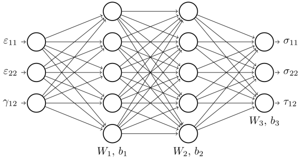
Offline
Physics
Data
Online

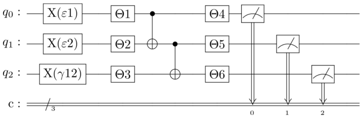
Path-dependent material behavior

Recurrent Neural Networks (RNN) for strain path dependency:

- Network includes latent (hidden) variables accounting for history dependency
- Fast surrogates for expensive models. Accurate as long as trained with enough data







[Giorgio Balducci, Poster session!]

[Boyang Chen's group (TUD)]

Offline

Physics

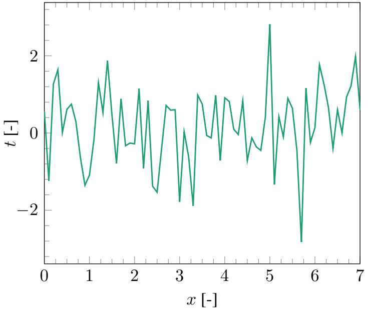
Data

Online

Active learning

Reducing sampling effort by only getting data when necessary:

- Expensive model response is unknown for most inputs $\Rightarrow$ epistemic uncertainty
- Bayesian machine learning can elegantly treat this problem



no correlation

$$p\left(\mathbf{t}\right) = \mathcal{N}\left(\mathbf{t}|\mathbf{0}, \sigma_{\mathrm{n}}^{2}\mathbf{I}\right)$$

Offline

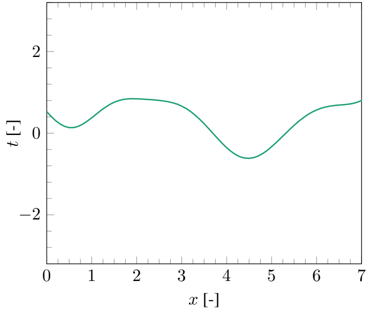
Physics Data

Online

Active learning

Reducing sampling effort by only getting data when necessary:

- Expensive model response is unknown for most inputs $\Rightarrow$ epistemic uncertainty
- Bayesian machine learning can elegantly treat this problem



$$p(\mathbf{t}) = \mathcal{N}(\mathbf{t}|\mathbf{0}, \mathbf{K})$$

$$K_{pq} = k(\mathbf{x}_p, \mathbf{x}_q) = \sigma_\mathrm{f}^2 \exp\left(-\frac{1}{2\ell^2}\left\|\mathbf{x}_p - \mathbf{x}_q\right\|^2\right)$$

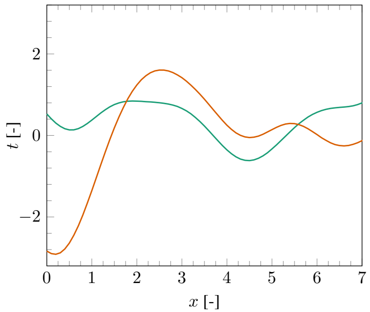
variance scaling

length scale

Active learning

Reducing sampling effort by only getting data when necessary:

- Expensive model response is unknown for most inputs $\Rightarrow$ epistemic uncertainty
- Bayesian machine learning can elegantly treat this problem



$$p\left(\mathbf{t}\right) = \mathcal{N}\left(\mathbf{t}|\mathbf{0},\mathbf{K}\right)$$

$$K_{pq} = k\left(\mathbf{x}_p,\mathbf{x}_q\right) = \sigma_{\mathrm{f}}^2 \exp\left(-\tfrac{1}{2\ell^2}\left\|\mathbf{x}_p-\mathbf{x}_q\right\|^2\right)$$

variance scaling length scale

Offline

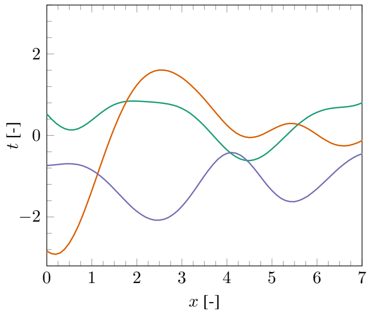
Physics Data

Online

Active learning

Reducing sampling effort by only getting data when necessary:

- Expensive model response is unknown for most inputs $\Rightarrow$ epistemic uncertainty
- Bayesian machine learning can elegantly treat this problem



$$p(\mathbf{t}) = \mathcal{N}(\mathbf{t}|\mathbf{0}, \mathbf{K})$$

$$K_{pq} = k(\mathbf{x}_p, \mathbf{x}_q) = \sigma_\mathrm{f}^2 \exp\left(-\frac{1}{2\ell^2} \|\mathbf{x}_p - \mathbf{x}_q\|^2\right)$$

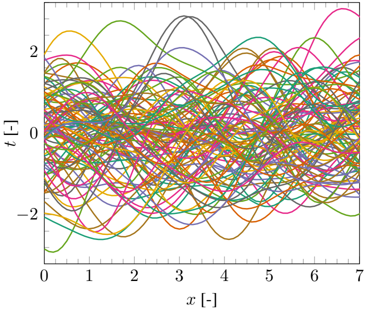
variance scaling

length scale

Active learning

Reducing sampling effort by only getting data when necessary:

- Expensive model response is unknown for most inputs $\Rightarrow$ epistemic uncertainty
- Bayesian machine learning can elegantly treat this problem



$$p(\mathbf{t}) = \mathcal{N}(\mathbf{t}|\mathbf{0}, \mathbf{K})$$

$$K_{pq} = k(\mathbf{x}_p, \mathbf{x}_q) = \sigma_{\mathrm{f}}^2 \exp\left(-\frac{1}{2\ell^2} \|\mathbf{x}_p - \mathbf{x}_q\|^2\right)$$

variance scaling   length scale

Offline

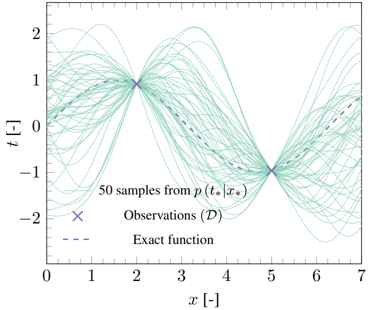
Physics Data

Online

Active learning

Reducing sampling effort by only getting data when necessary:

- Expensive model response is unknown for most inputs $\Rightarrow$ epistemic uncertainty
- Bayesian machine learning can elegantly treat this problem



data likelihood

prior

$$p\left(\mathbf{t}_*|\mathbf{t}_\mathrm{o}\right) = \frac{p(\mathbf{t}_\mathrm{o}|\mathbf{t}_*)p(\mathbf{t}_*)}{p(\mathbf{t}_\mathrm{o})}$$

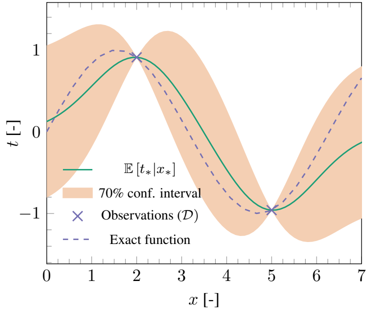
posterior

model evidence

Active learning

Reducing sampling effort by only getting data when necessary:

- Expensive model response is unknown for most inputs $\Rightarrow$ epistemic uncertainty
- Bayesian machine learning can elegantly treat this problem



data likelihood

prior

$$p\left(\mathbf{t}_* | \mathbf{t}_\mathrm{o}\right) = \frac{p(\mathbf{t}_\mathrm{o}|\mathbf{t}_*)p(\mathbf{t}_*)}{p(\mathbf{t}_\mathrm{o})}$$

posterior

model evidence

Offline

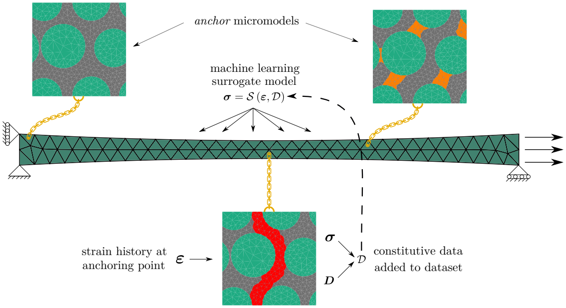
Physics Data

Online

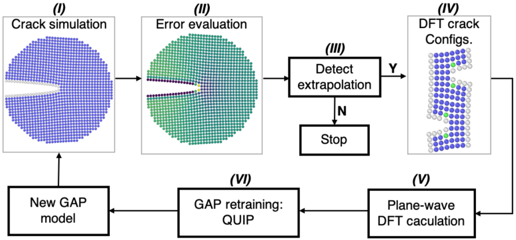
Active learning

Reducing sampling effort by only getting data when necessary:

- Expensive model response is unknown for most inputs ⇒ epistemic uncertainty
- Bayesian machine learning can elegantly treat this problem



[Rocha et al (2021), JCPX 9:100083]



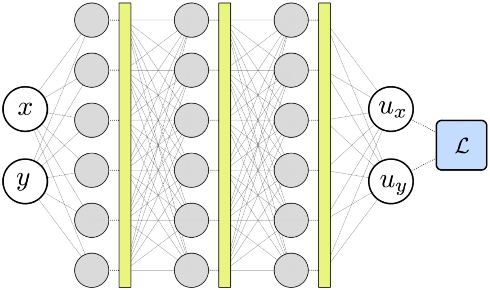
[Zhang et al (2022), arXiv:2208.05912v1]

[Francesco Maresca's group (RUG)]

Physics-informed neural networks

Neural networks as PDE solvers:

- BC/IC values treated as conventional observations
- Residual of strong-form PDE at collocation points added to loss function



[Haghighat et al (2020), arXiv:2003.02751v2]

$$\sigma_{ij,j} + f_i = 0$$

$$\sigma_{ij} = \lambda \delta_{ij} \varepsilon_{kk} + 2\mu \varepsilon_{ij}$$

$$\varepsilon_{ij} = \frac{1}{2}\left(u_{i,j} + u_{j,i}\right)$$

$$\mathcal{L} = |\sigma_{ij,j} + f_i|_\Omega$$

Col. point
PDE loss

Offline

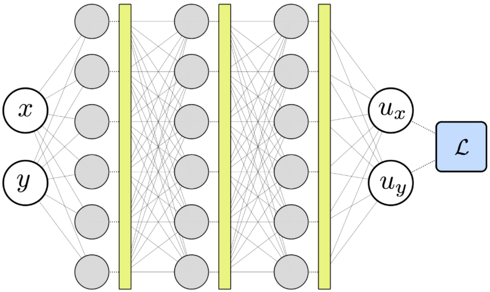
Physics Data

Online

Physics-informed neural networks

Neural networks as PDE solvers:

- BC/IC values treated as conventional observations
- Residual of strong-form PDE at collocation points added to loss function



[Haghighat et al (2020), arXiv:2003.02751v2]

$$\sigma_{ij,j} + f_i = 0$$

$$\sigma_{ij} = \lambda \delta_{ij} \varepsilon_{kk} + 2\mu \varepsilon_{ij}$$

$$\varepsilon_{ij} = \frac{1}{2}\left(u_{i,j} + u_{j,i}\right)$$

$$\mathcal{L} = |\sigma_{ij,j} + f_i|_\Omega + |u - u^*|_{\Gamma_u}$$

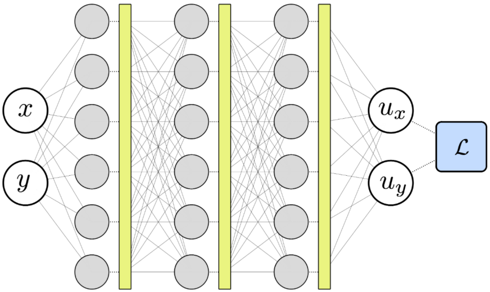
Col. point PDE loss

Dirichlet loss

Physics-informed neural networks

Neural networks as PDE solvers:

- BC/IC values treated as conventional observations
- Residual of strong-form PDE at collocation points added to loss function



[Haghighat et al (2020), arXiv:2003.02751v2]

$$\sigma_{ij,j} + f_i = 0$$

$$\sigma_{ij} = \lambda \delta_{ij} \varepsilon_{kk} + 2\mu \varepsilon_{ij}$$

$$\varepsilon_{ij} = \frac{1}{2}\left(u_{i,j} + u_{j,i}\right)$$

$$\mathcal{L} = |\sigma_{ij,j} + f_i|_{\Omega} + |u - u^*|_{\Gamma_u} + |\sigma_{ij} - \sigma^*_{ij}|_{\Gamma_\sigma}$$

Col. point PDE loss | Dirichlet loss | Neumann loss

Offline

Physics Data

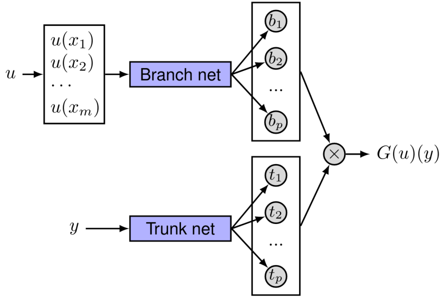
Online

Operator Networks

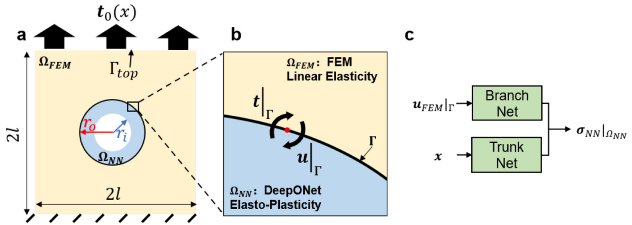
Learning functional operators from data:

- Beneficial bias through architecture split $\Rightarrow$ better generalization

Deep Operator Network (DeepONet)



[Lu et al (2020), arXiv:1910.03193v3]



[Yin et al (2022), CMAME, In press]

Offline

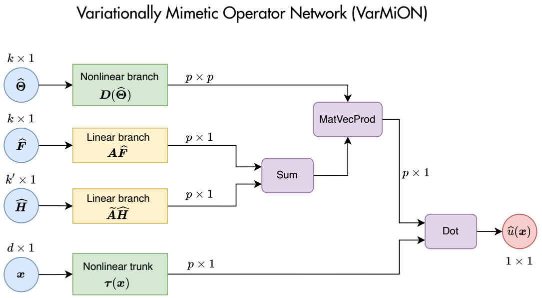
Physics Data

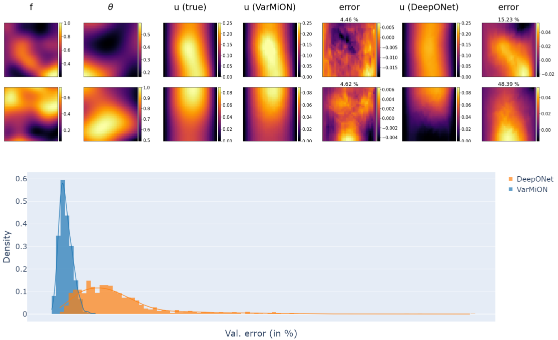
Online

Operator Networks

Learning functional operators from data:

- Beneficial bias through architecture split $\Rightarrow$ better generalization





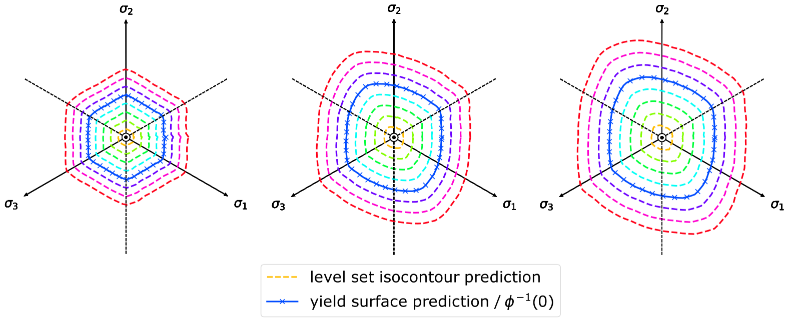
[Patel et al (2022), arXiv:2209:12871]

[Michael Abdelmalik's group (TU/e)]

Hybrid models

Physical bias through constitutive assumptions:

- Machine learning for individual model components
- Embedding complete physical models in ML framework



[Vlassis and Sun (2021), CMAME 377:113695]

Offline

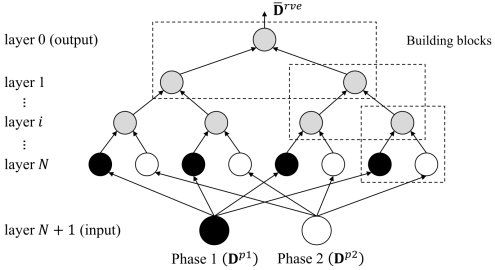
Physics Data

Online

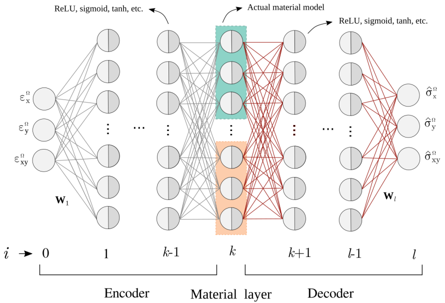
Hybrid models

Physical bias through constitutive assumptions:

- Machine learning for individual model components
- Embedding complete physical models in ML framework



[Liu et al (2019), CMAME 345:1138-1168]



[Maia et al (2022), arXiv:2209.07320v1]

Offline

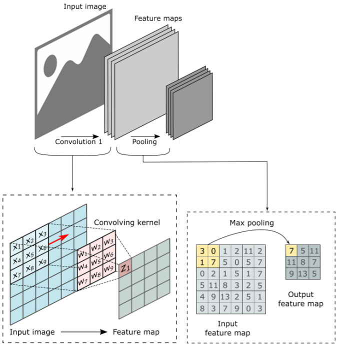
Physics Data

Online

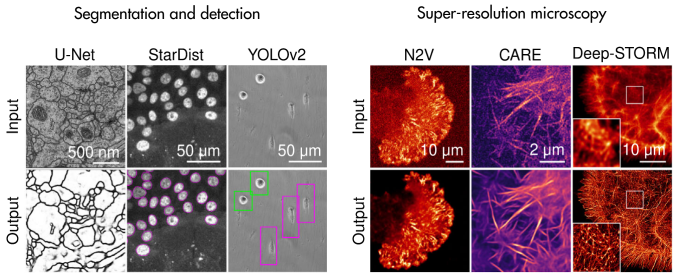
Computer vision

Extracting knowledge from images in creative ways:

- Augmenting microscopy experiments
- Reduced-order model selection



[Adapted from: Alzubaidi et al (2022), Rock Mech Rock Eng 55:3719-3734]



[Adapted from: Von Chamier et al (2021), Nature Comm 12:2276]

Offline

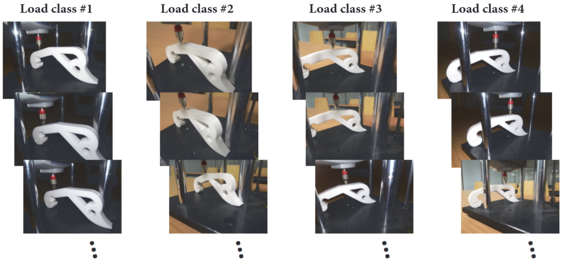
Physics Data

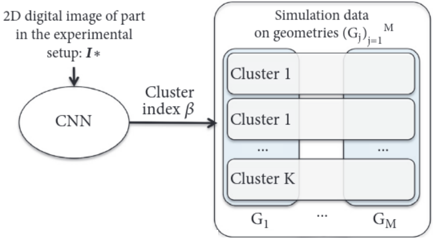
Online

Computer vision

Extracting knowledge from images in creative ways:

- Augmenting microscopy experiments
- Reduced-order model selection





[N'Guyen et al (2018), Complexity 2018:10]

Offline

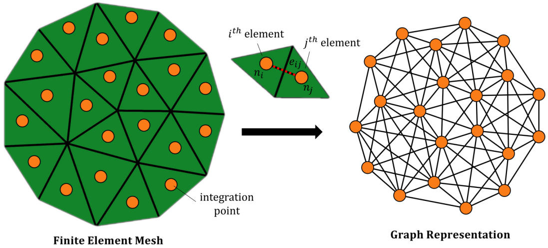
Physics Data

Online

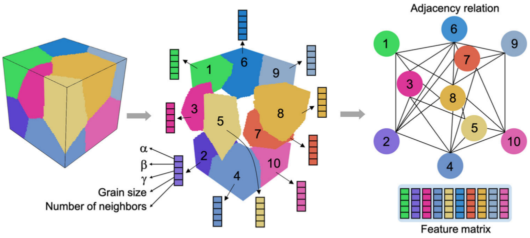
Deep learning on graphs

Inductive bias coming from geometry:

- Sparse network connectivity following a priori geometric assumptions
- Information spreads throughout the domain through message passing



[Vlassis and Sun (2022), arXiv:2208.00246v1]



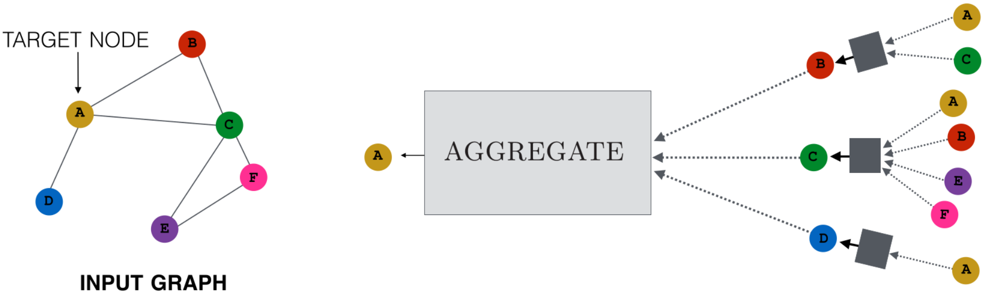
[Dai et al (2021), npj Comput Mat 7:103]

Offline
Physics Data
Online

Deep learning on graphs

Inductive bias coming from geometry:

- Sparse network connectivity following a priori geometric assumptions
- Information spreads throughout the domain through message passing



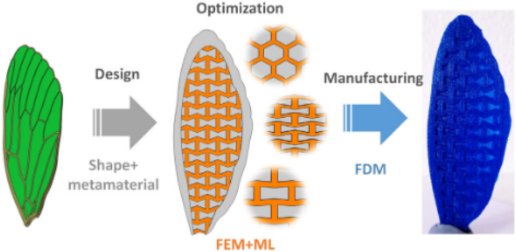
[Hamilton (2020), Graph Representation Learning, Lecture notes]

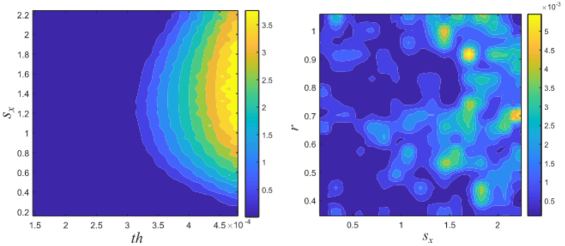
Offline
Physics
Data
Online

Optimization and discovery with ML

Learning architectures for material design:

- Design exploration, Bayesian optimization
- Generative machine learning





[Zhilyaev et al (2022), Materials and Design 218:110709]

[Anastasiia Krushynska's group (RUG)]

Offline

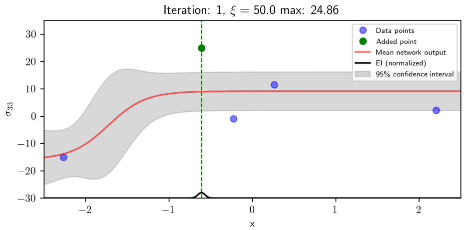
Physics Data

Online

Optimization and discovery with ML

Learning architectures for material design:

- Design exploration, Bayesian optimization
- Generative machine learning



[J Storm (2021), MSc Thesis]

Offline

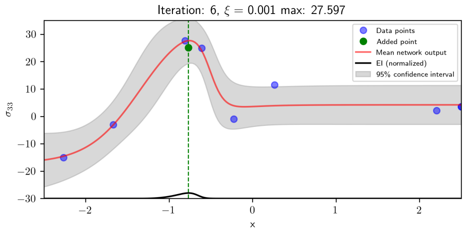
Physics Data

Online

Optimization and discovery with ML

Learning architectures for material design:

- Design exploration, Bayesian optimization
- Generative machine learning



[J Storm (2021), MSc Thesis]

Offline

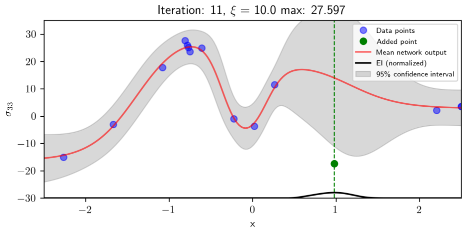
Physics Data

Online

Optimization and discovery with ML

Learning architectures for material design:

- Design exploration, Bayesian optimization
- Generative machine learning



[J Storm (2021), MSc Thesis]

Offline

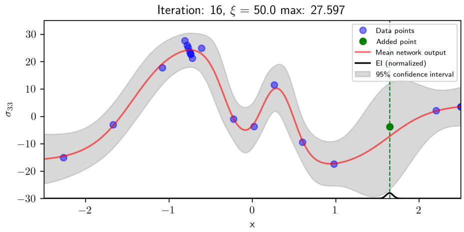
Physics Data

Online

Optimization and discovery with ML

Learning architectures for material design:

- Design exploration, Bayesian optimization
- Generative machine learning



[J Storm (2021), MSc Thesis]

Offline

Physics

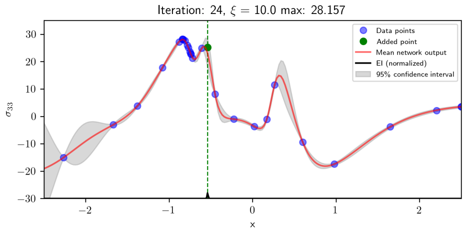
Data

Online

Optimization and discovery with ML

Learning architectures for material design:

- Design exploration, Bayesian optimization
- Generative machine learning



[J Storm (2021), MSc Thesis]

Offline

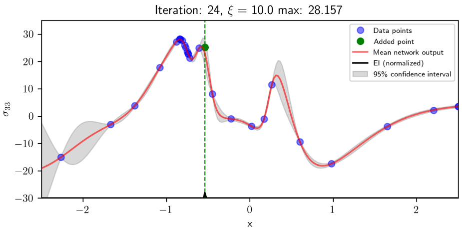
Physics Data

Online

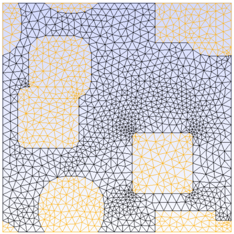
Optimization and discovery with ML

Learning architectures for material design:

- Design exploration, Bayesian optimization
- Generative machine learning



[J Storm (2021), MSc Thesis]

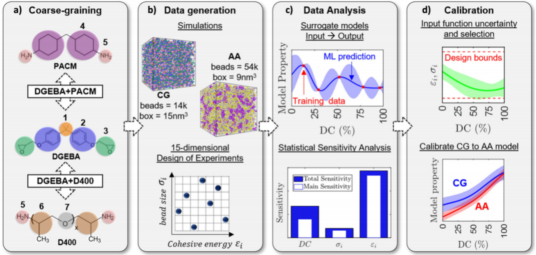


Offline
Physics
Data
Online

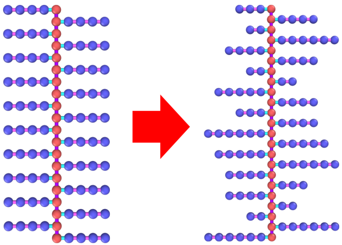
Optimization and discovery with ML

Learning architectures for material design:

- Design exploration, Bayesian optimization
- Generative machine learning



Bayesian optimization for polydispersity in branched polymers

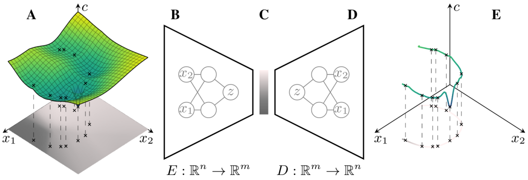


[Andrea Giuntoli's group (RUG)]

Optimization and discovery with ML

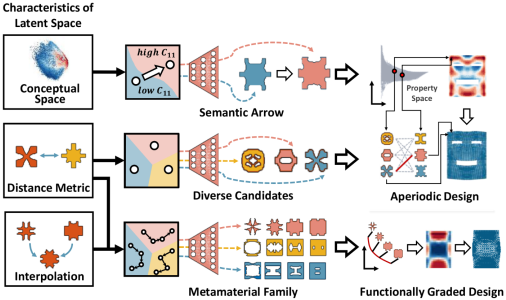
Learning architectures for material design:

- Design exploration, Bayesian optimization
- Generative machine learning



[Schumann and Aragón (2022), arXiv:2110.14985v1]

[Alejandro Aragón's group (TUD)]



[Wang et al (2020), CMAME 372:113377]

Introducing our workshop speakers

TU Eindhoven: Bas Kessels

- ML-based parameter updating in nonlinear dynamics

Twente University: Retief Lubbe

- Bayesian inference of granular mesostructures

University of Groningen: Lei Zhang

- Active learning for atomistic models

TU Delft: Prakash Thakolkaran

- Learning hyperelasticity without stress data

We hope to see you at the workshop!

A nervous young researcher about to present their work, by Frans Hals