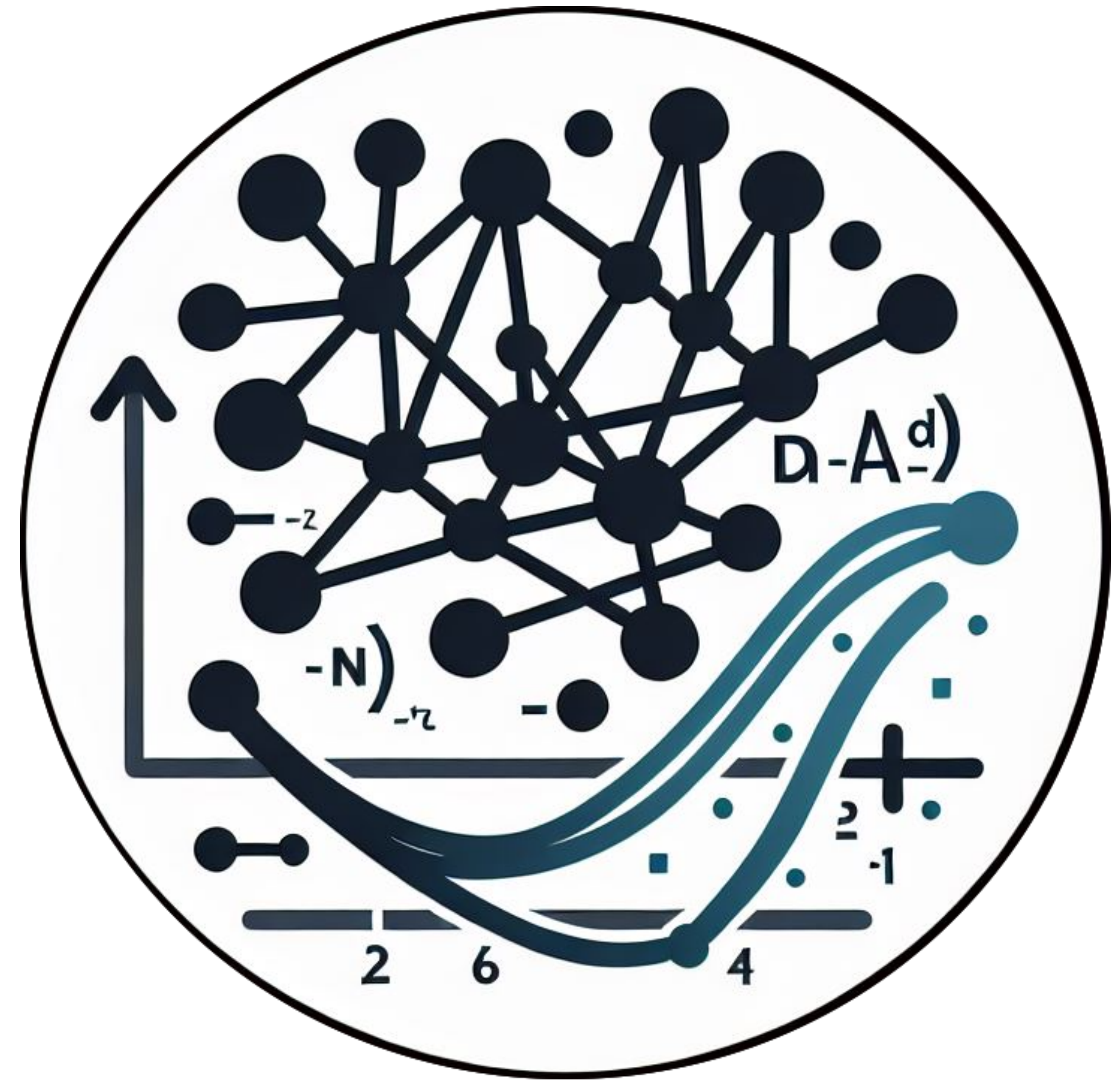


Review

Deep Learning for Engineers

Andrew Ning

aning@byu.edu



Logistics

First a few random topics

Supervised. has labels. regression and classification.

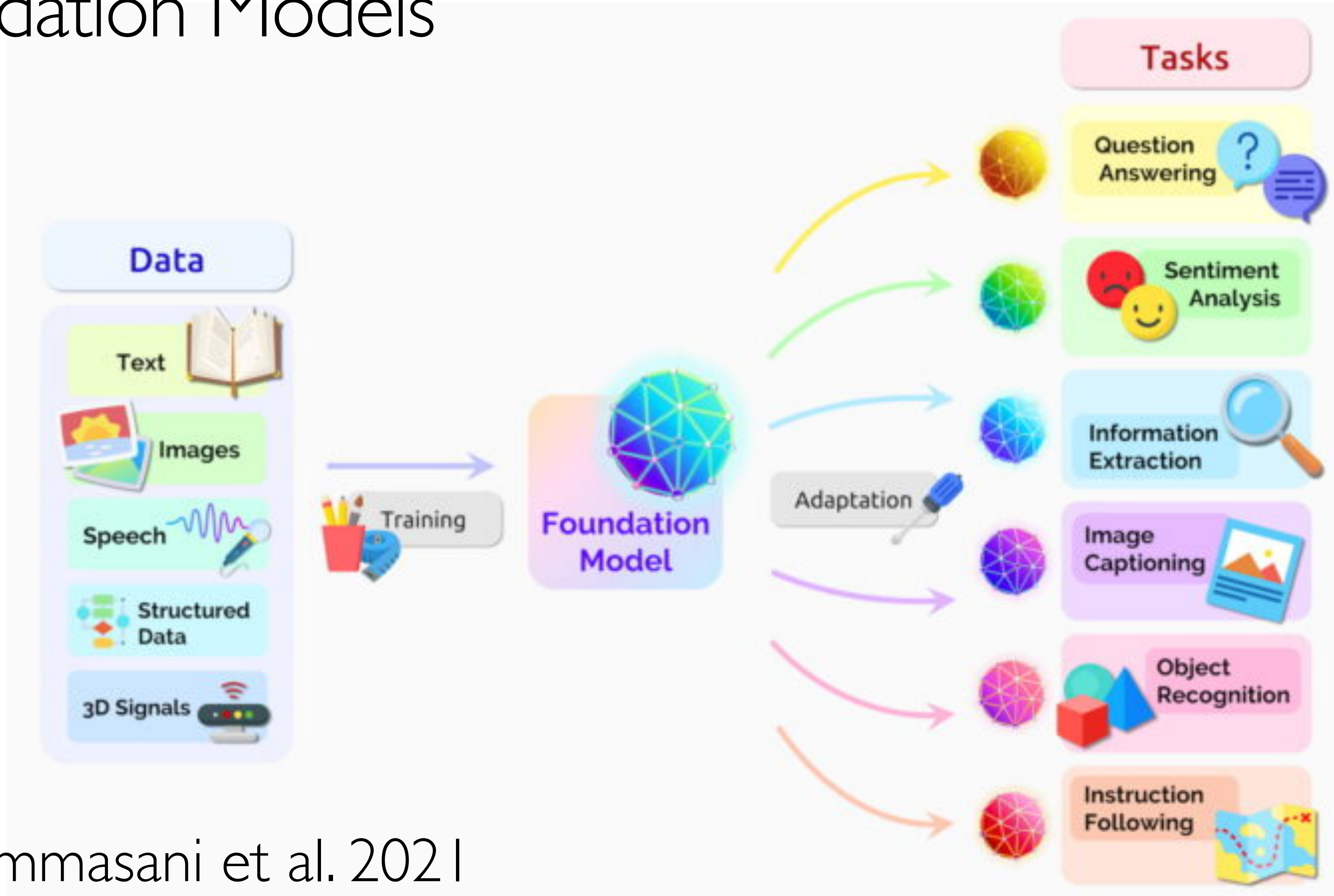
Unsupervised. no labels. clustering, pattern discovery.

Reinforcement. interact with environment. maximize rewards over time.



Hyperparameter Optimization

Foundation Models



Bommasani et al. 2021

Training compute of notable machine learning models by domain, 2012–23

Source: Epoch, 2023 | Chart: 2024 AI Index report

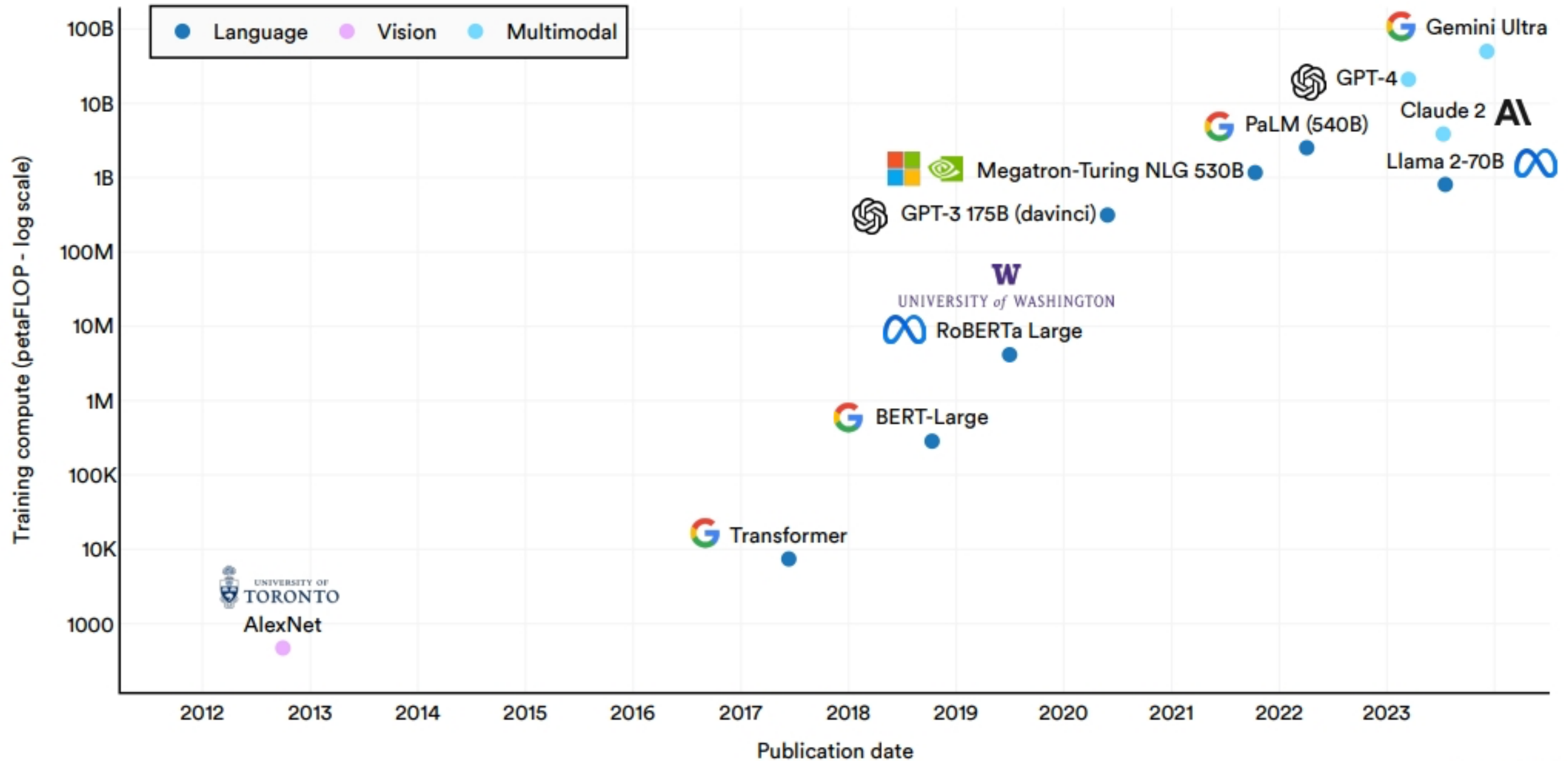
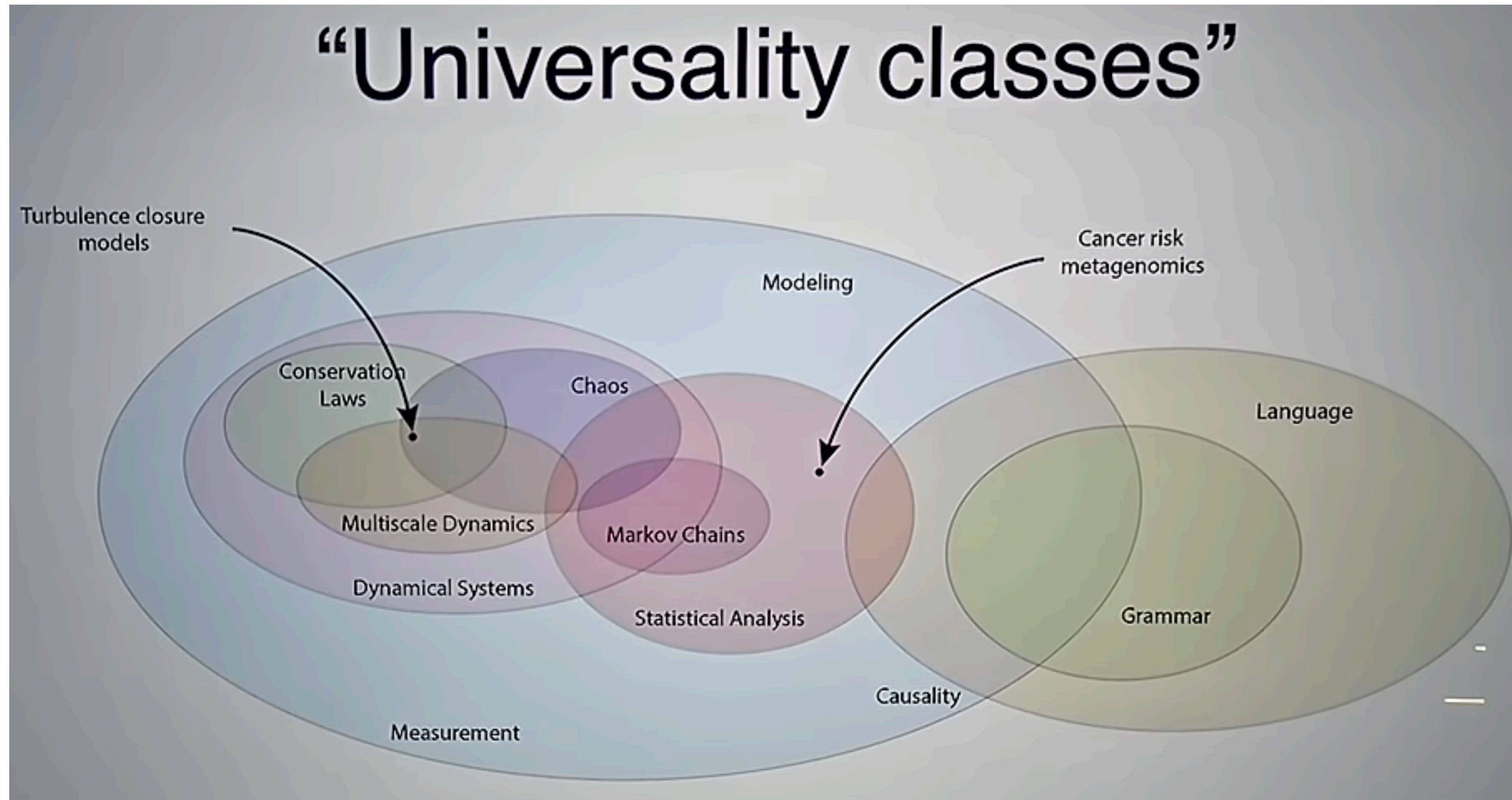
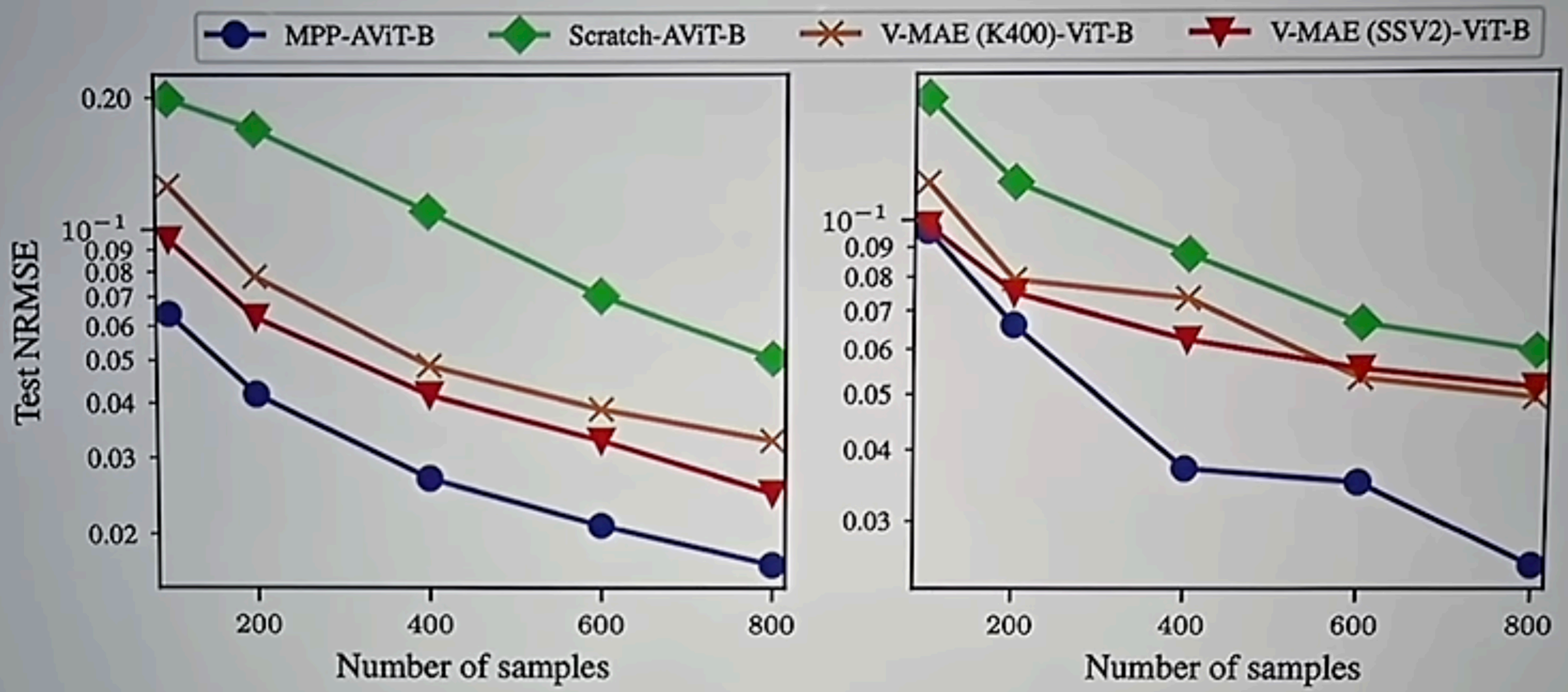


Figure 1.3.7

Engineering Foundation Models



Miles Cranmer - The Next Great Scientific Theory is Hiding Inside a Neural Network

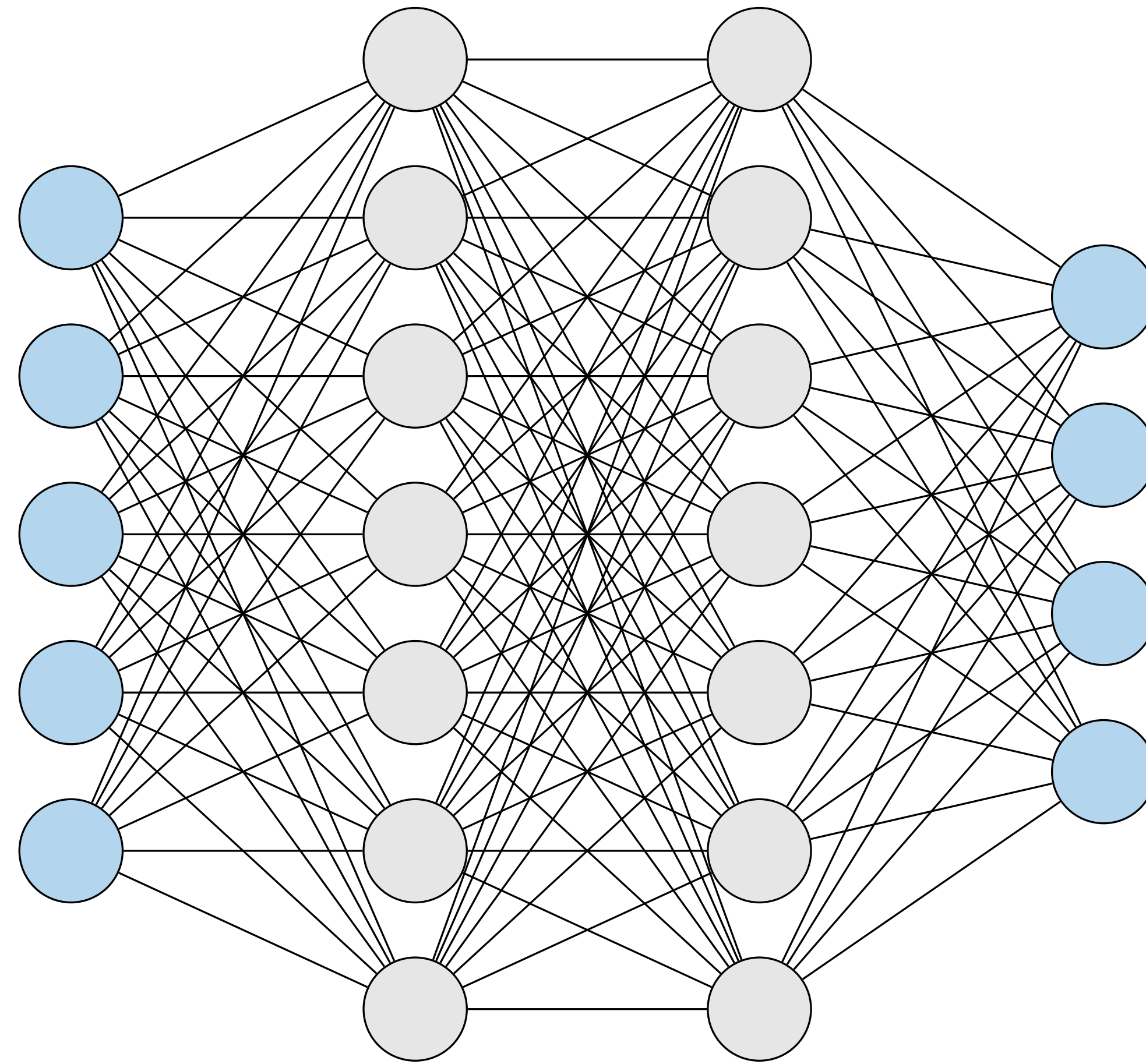


<https://github.com/PolymathicAI/walrus>

Review

https://flow.byu.edu/me595r/schedule/dl_architecture_guide.html

MLP

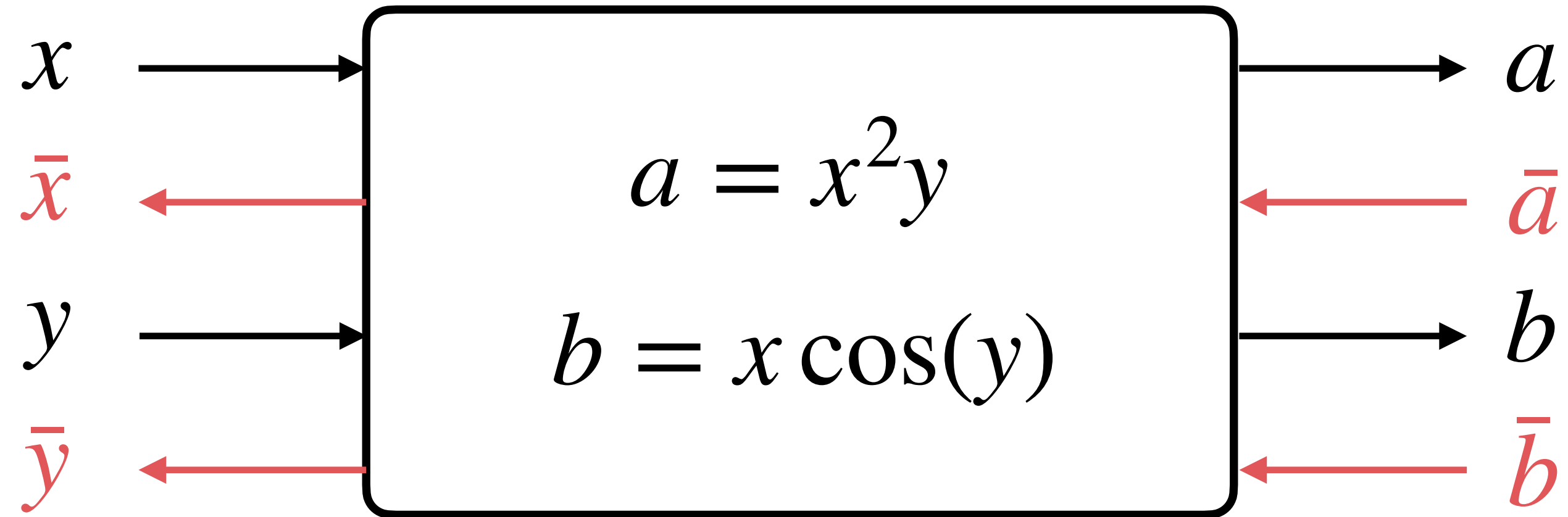


Input layer

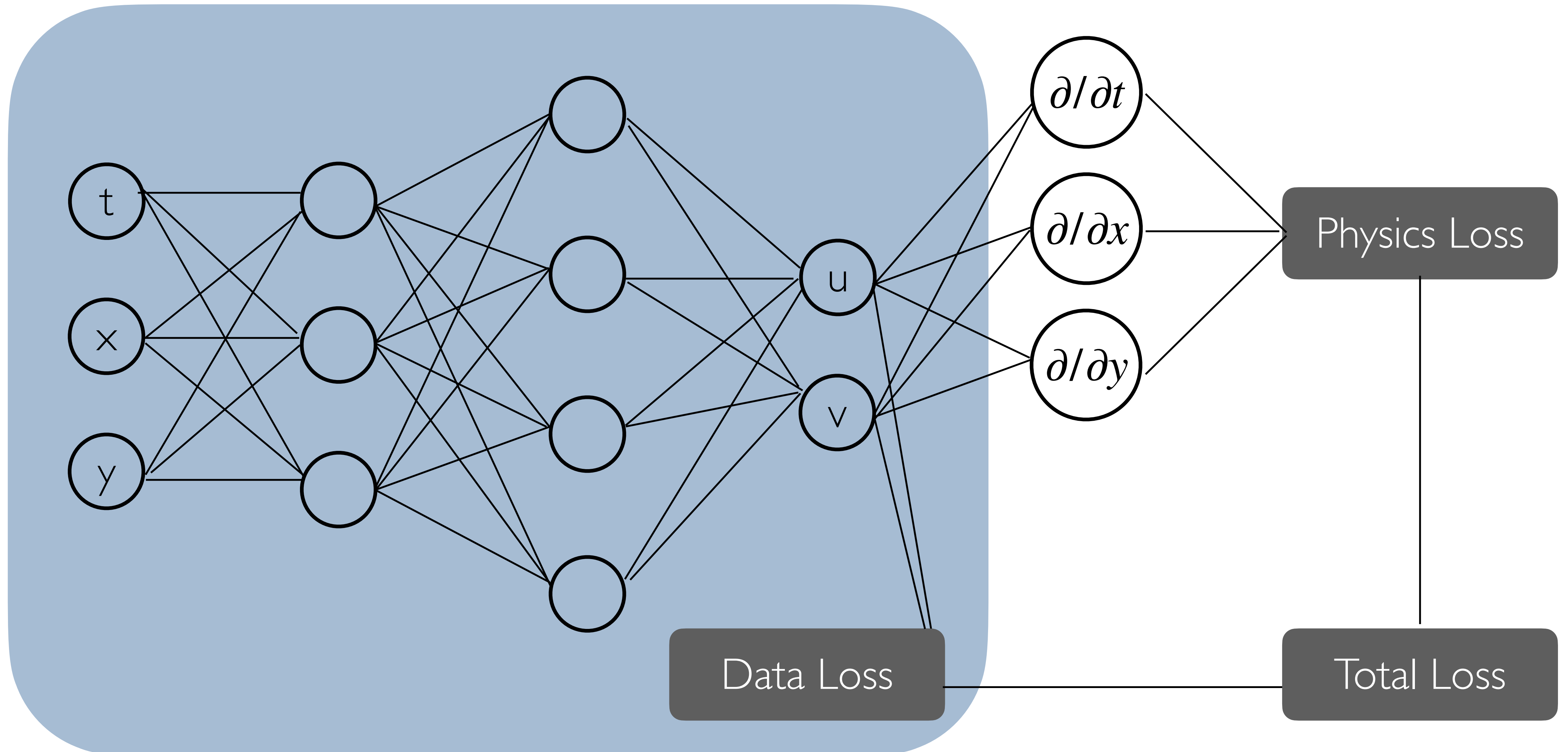
Hidden layers

Output layer

Backprop



PINN



Inverse PINN

$$L = \frac{1}{m} \sum_{i=1}^m (\hat{y}_i - y_i)^2 + \gamma \frac{1}{n} \sum_{j=1}^n \left(m \frac{\partial^2 \hat{y}}{\partial t^2} + \mu \frac{\partial \hat{y}}{\partial t} + k \hat{y} \right)_j^2$$

Optimizers

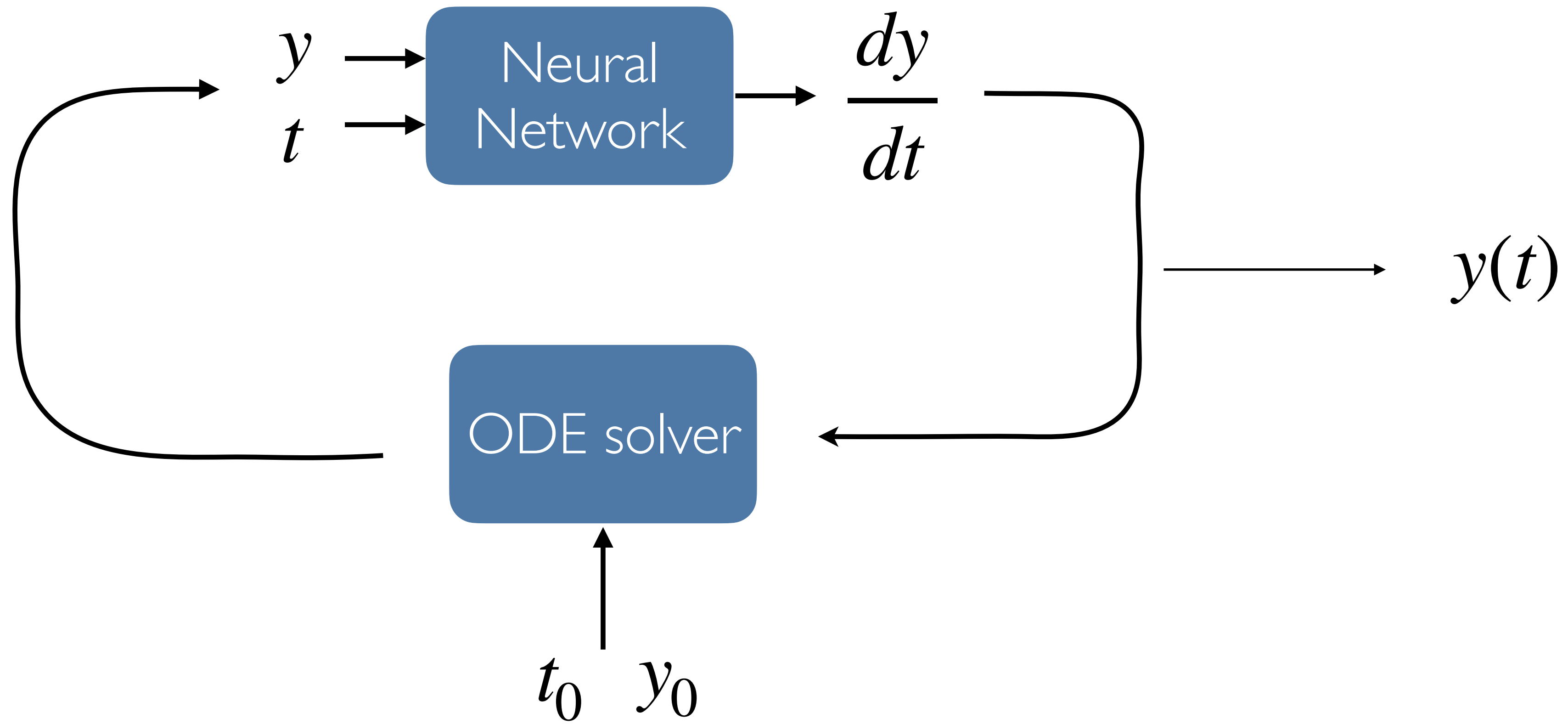
SGD

Momentum: smooth out gradient (direction) with exponential moving average

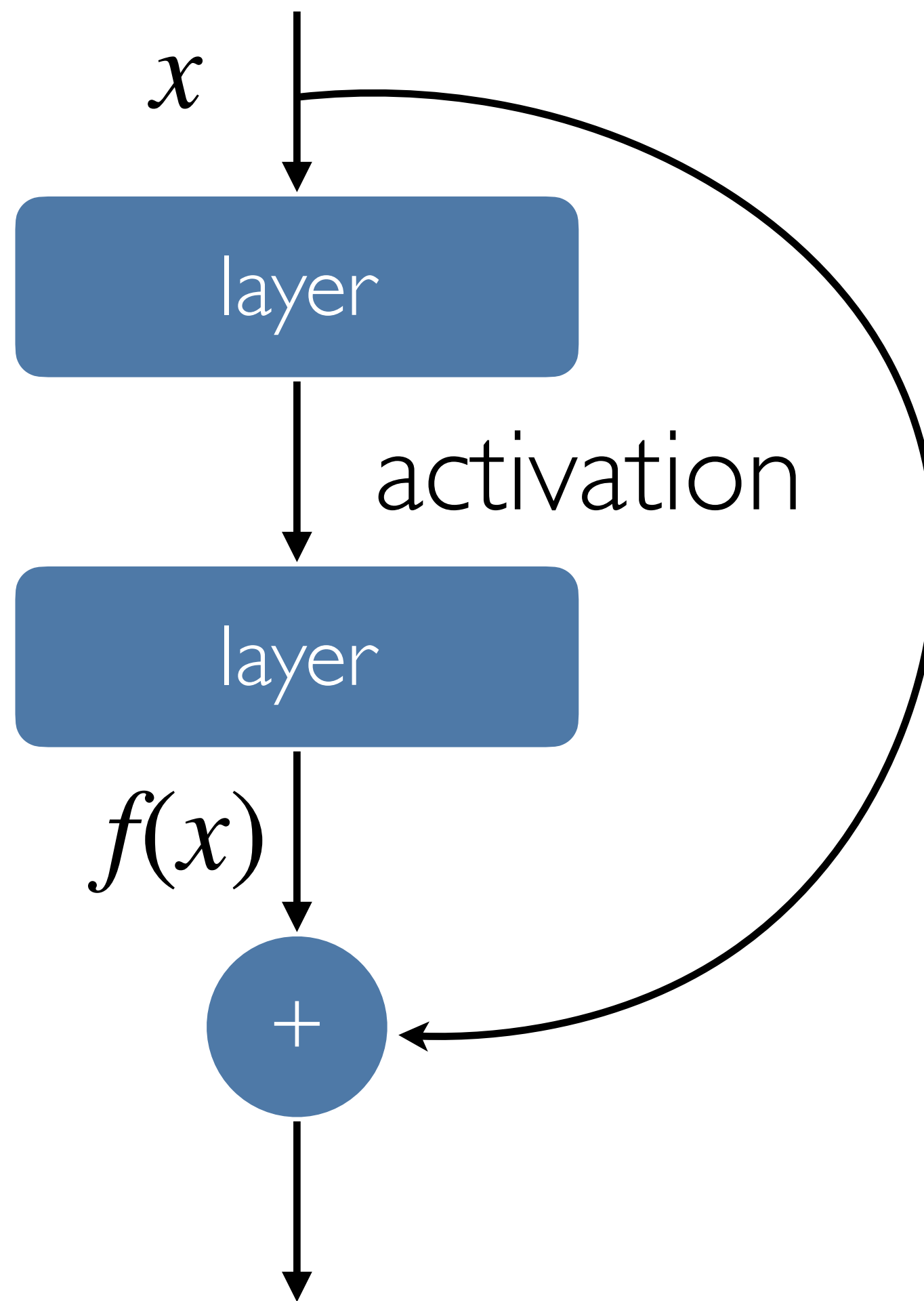
RMSprop: modify step size using EMA of squared gradient (intuition is that if gradient is small we should take larger steps and if gradient is large we should take smaller steps)

Adam: do both

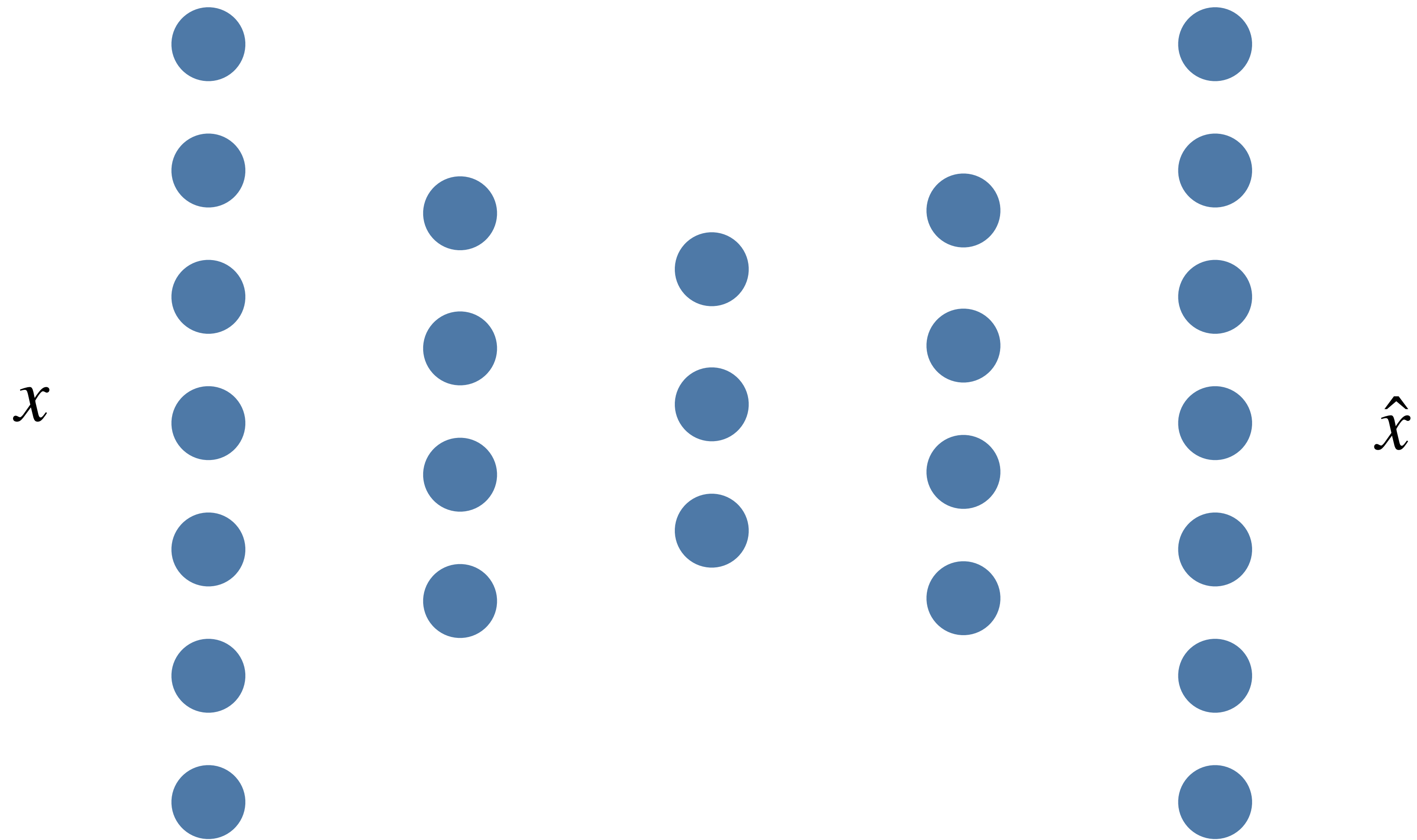
NeuralODE



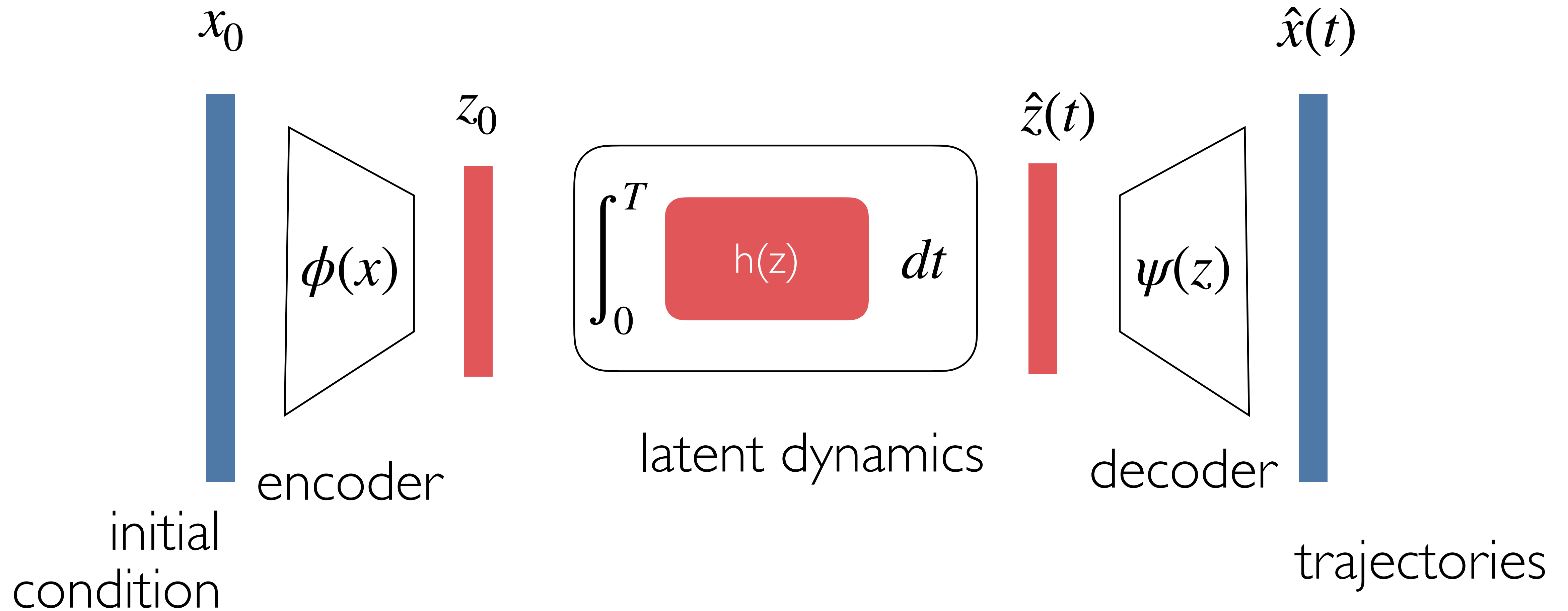
ResNet



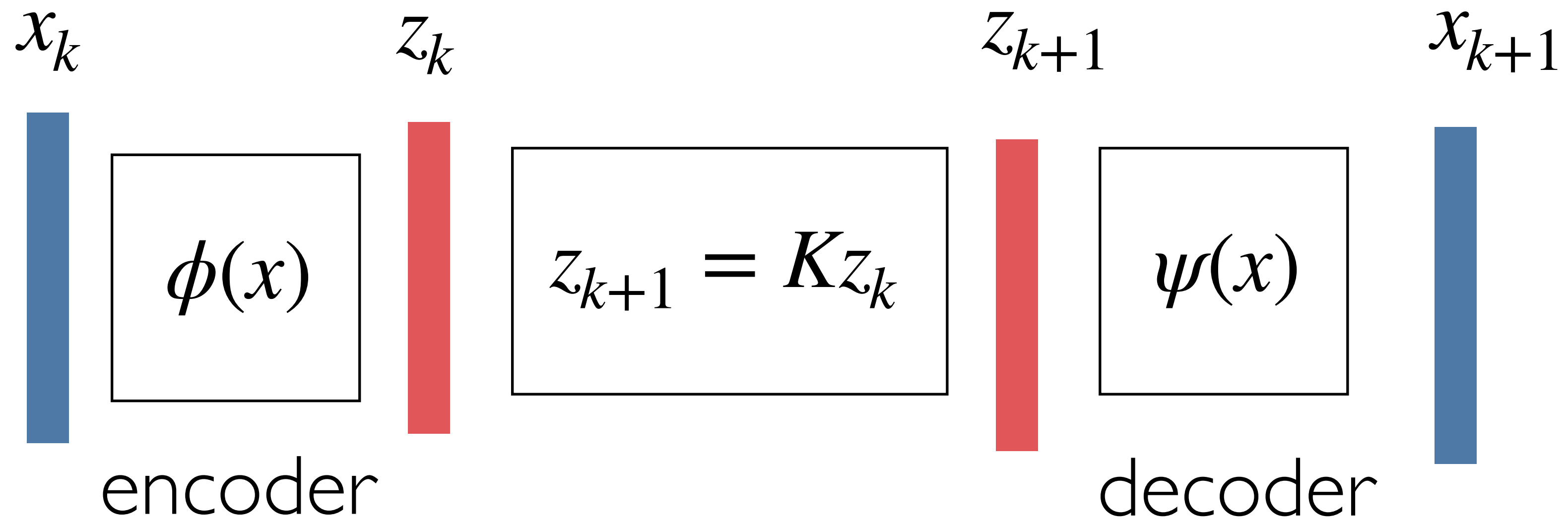
Autoencoders



Neural ODE with autoencoder



Neural Koopman



Regularization

more data

more physics

early stopping

weight regularization

dropout

ensemble averaging

Convolutional Neural Nets

0	1	2
2	2	0
0	1	2

3 ₀	3 ₁	2 ₂	1	0
0 ₂	0 ₂	1 ₀	3	1
3 ₀	1 ₁	2 ₂	2	3
2	0	0	2	2
2	0	0	0	1

12.0	12.0	17.0
10.0	17.0	19.0
9.0	6.0	14.0

3	3 ₀	2 ₁	1 ₂	0
0	0 ₂	1 ₂	3 ₀	1
3	1 ₀	2 ₁	2 ₂	3
2	0	0	2	2
2	0	0	0	1

12.0	12.0	17.0
10.0	17.0	19.0
9.0	6.0	14.0

3	3	2 ₀	1 ₁	0 ₂
0	0	1 ₂	3 ₂	1 ₀
3	1	2 ₀	2 ₁	3 ₂
2	0	0	2	2
2	0	0	0	1

12.0	12.0	17.0
10.0	17.0	19.0
9.0	6.0	14.0

3	3	2	1	0
0 ₀	0 ₁	1 ₂	3	1
3 ₂	1 ₂	2 ₀	2	3
2 ₀	0 ₁	0 ₂	2	2
2	0	0	0	1

12.0	12.0	17.0
10.0	17.0	19.0
9.0	6.0	14.0

3	3	2	1	0
0	0 ₀	1 ₁	3 ₂	1
3	1 ₂	2 ₂	2 ₀	3
2	0 ₀	0 ₁	2 ₂	2
2	0	0	0	1

12.0	12.0	17.0
10.0	17.0	19.0
9.0	6.0	14.0

3	3	2	1	0
0	0	1 ₀	3 ₁	1 ₂
3	1	2 ₂	2 ₂	3 ₀
2	0	0 ₀	2 ₁	2 ₂
2	0	0	0	1

12.0	12.0	17.0
10.0	17.0	19.0
9.0	6.0	14.0

3	3	2	1	0
0	0	1	3	1
3 ₀	1 ₁	2 ₂	2	3
2 ₂	0 ₂	0 ₀	2	2
2 ₀	0 ₁	0 ₂	0	1

12.0	12.0	17.0
10.0	17.0	19.0
9.0	6.0	14.0

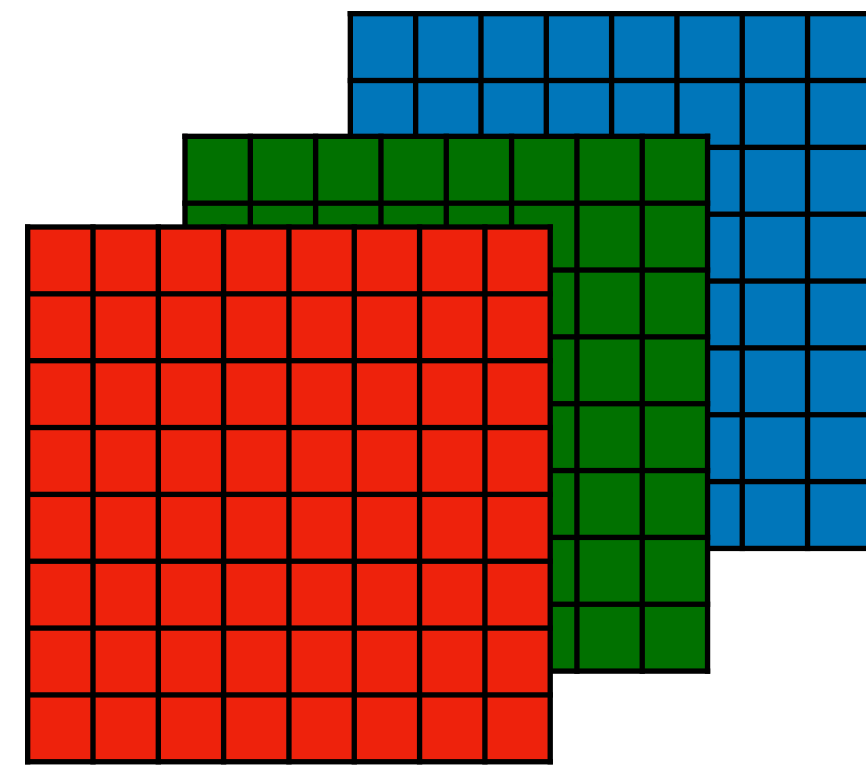
3	3	2	1	0
0	0	1	3	1
3	1 ₀	2 ₁	2 ₂	3
2	0 ₂	0 ₂	2 ₀	2
2	0 ₀	0 ₁	0 ₂	1

12.0	12.0	17.0
10.0	17.0	19.0
9.0	6.0	14.0

3	3	2	1	0
0	0	1	3	1
3	1	2 ₀	2 ₁	3 ₂
2	0	0 ₂	2 ₂	2 ₀
2	0	0 ₀	0 ₁	1 ₂

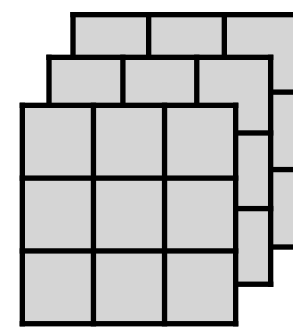
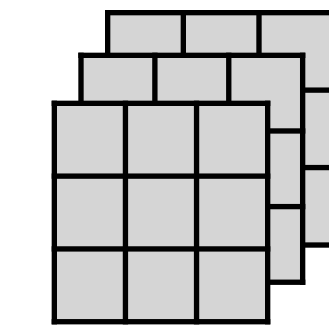
12.0	12.0	17.0
10.0	17.0	19.0
9.0	6.0	14.0

Multiple Inputs and Multiple Outputs



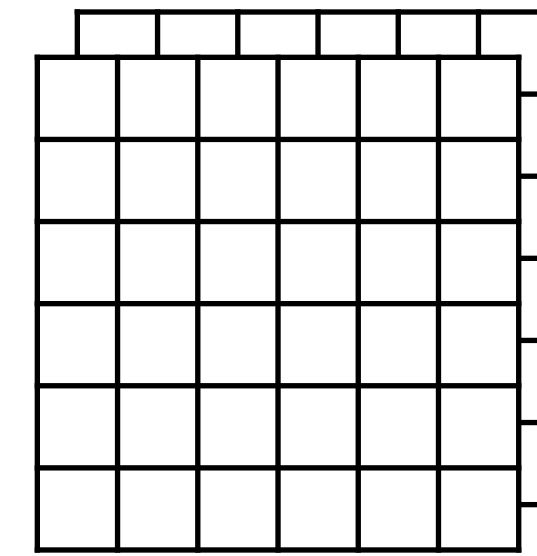
$8 \times 8 \times 3$

in channels



$3 \times 3 \times 3 \times 2$

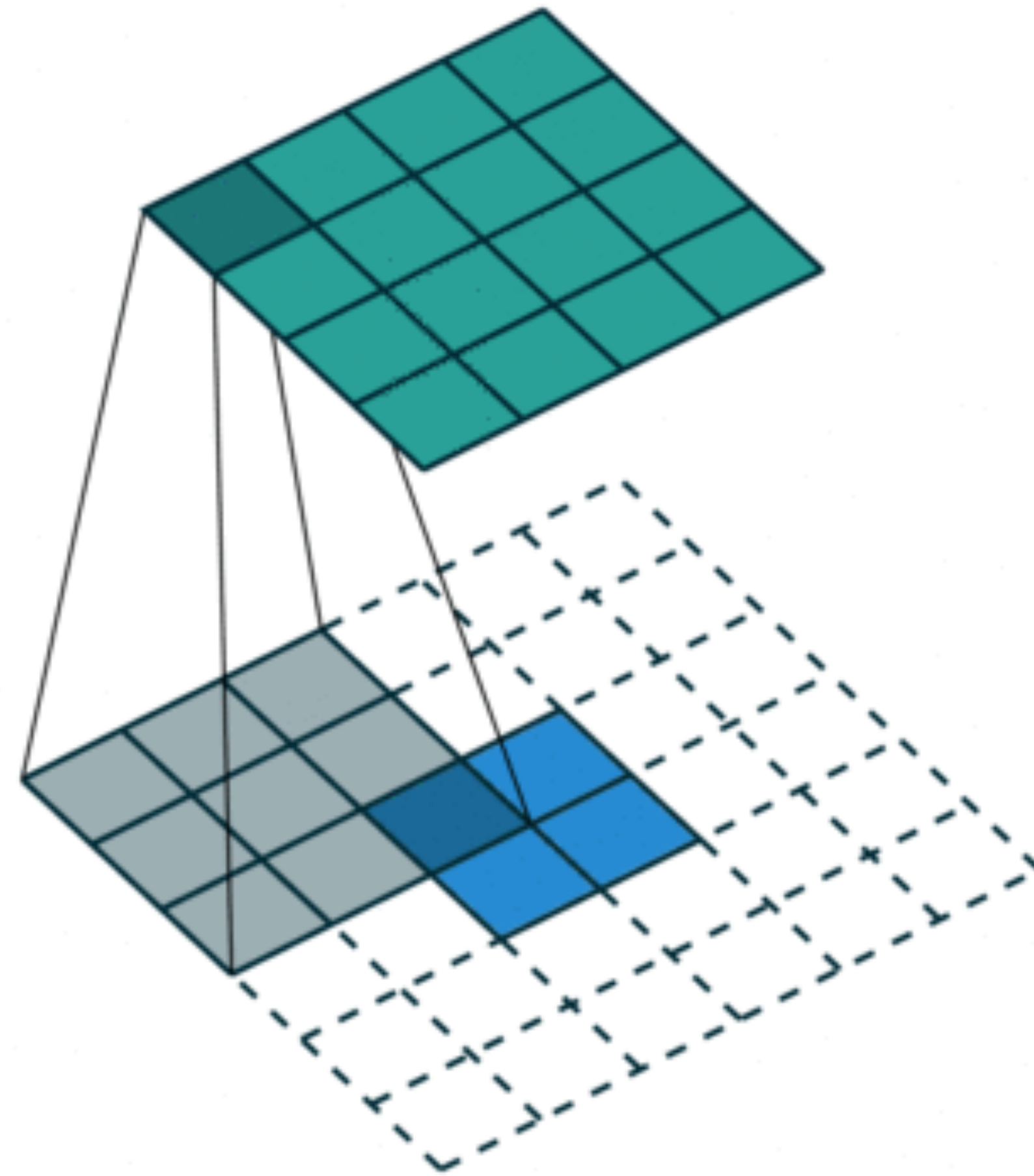
kernel size



$6 \times 6 \times 2$

out channels

Transpose Convolution

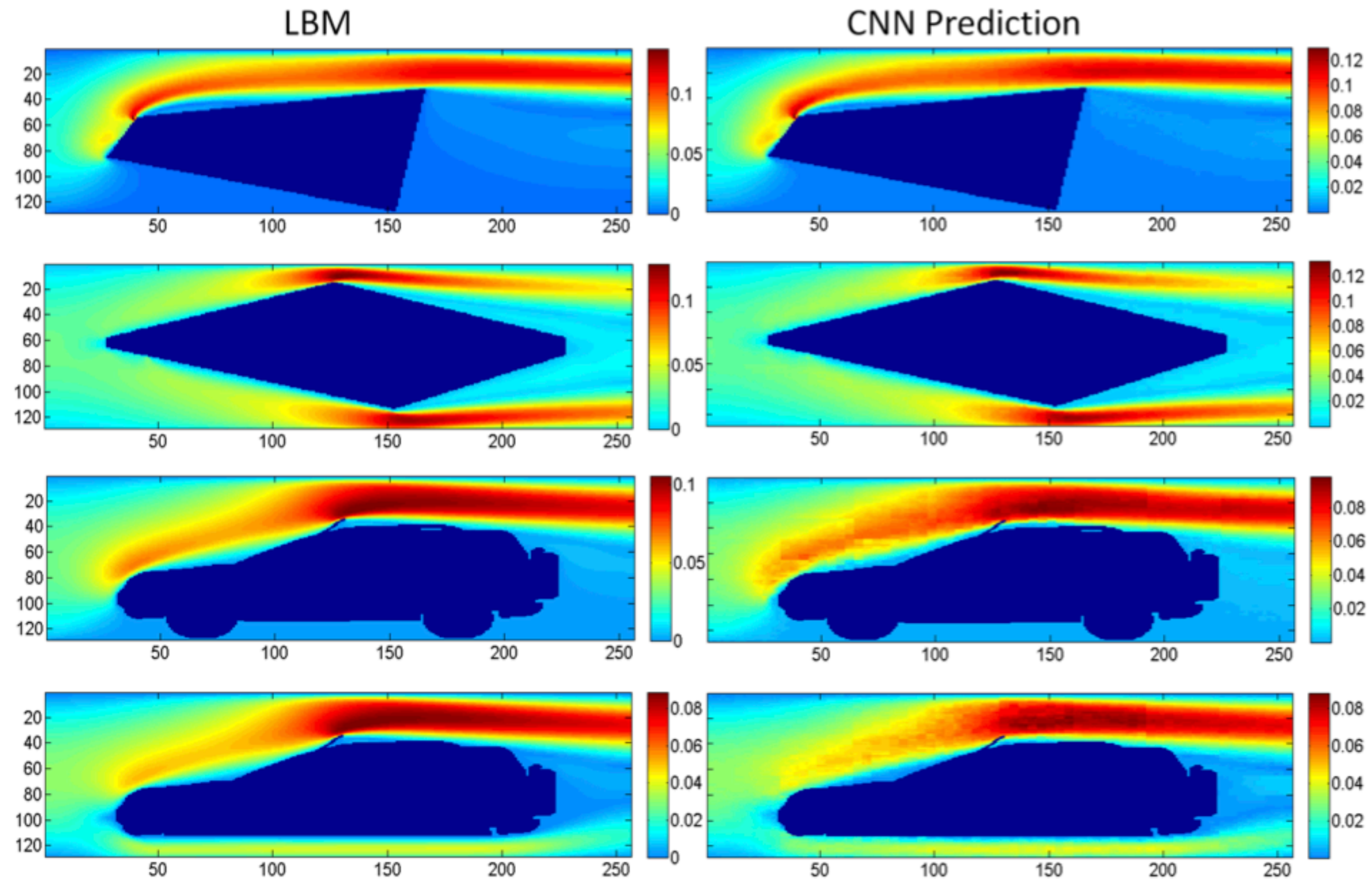


output

kernel

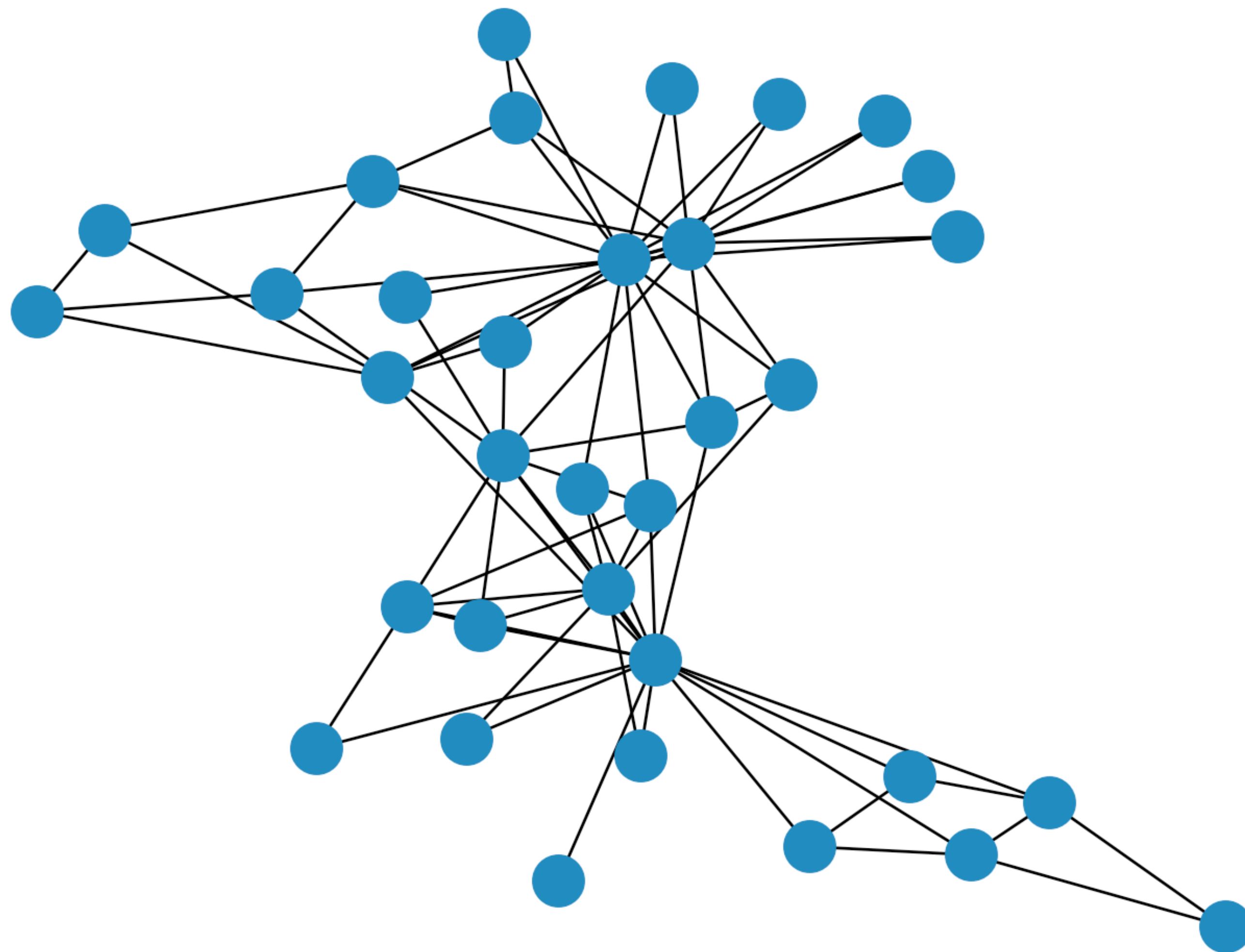
input

Spatial Regression

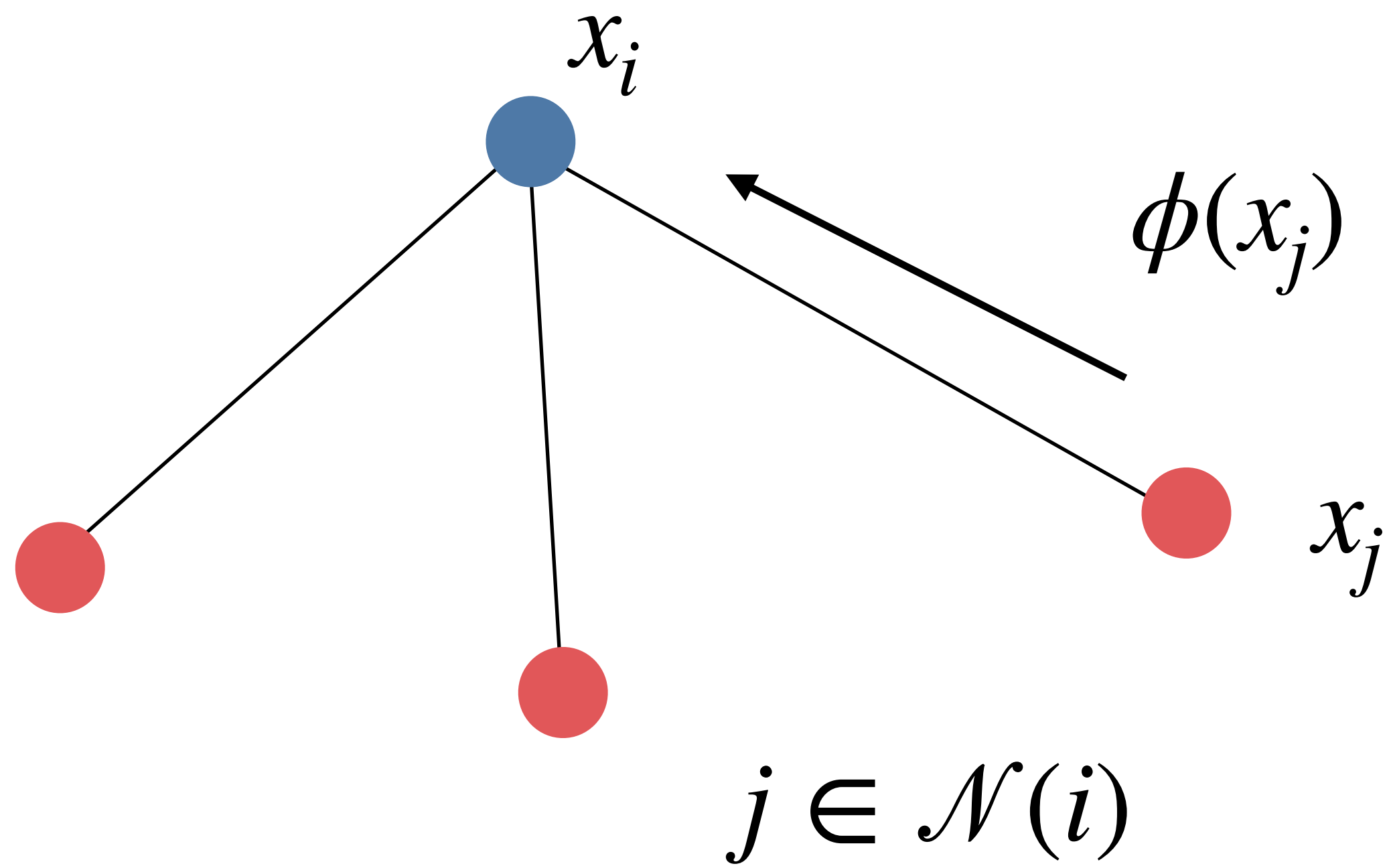


Convolutional Neural Networks for Steady Flow Approximation, Xiaoxiao Guo, Wei Li, Francesco Iorio

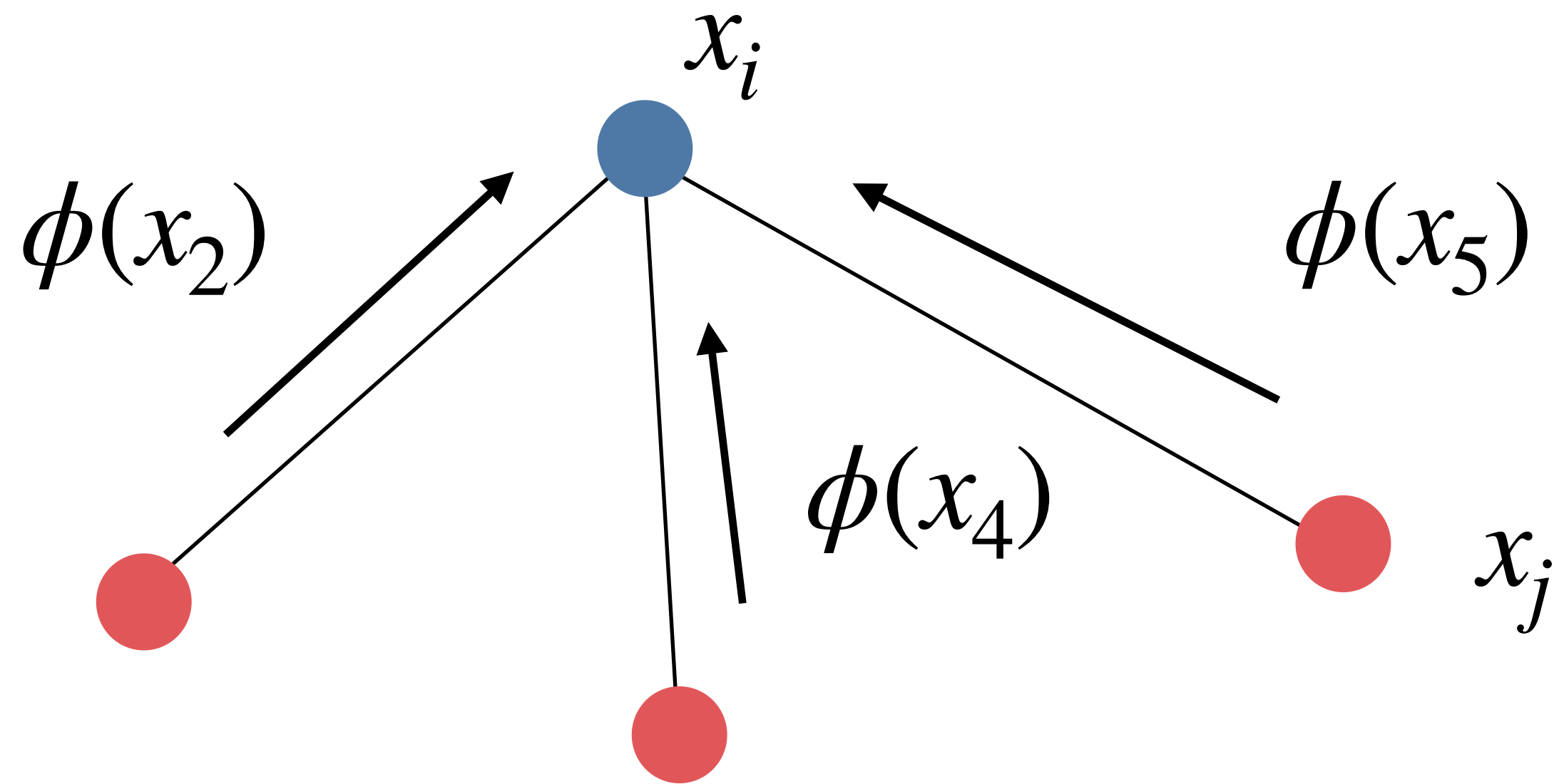
Graph Neural Nets



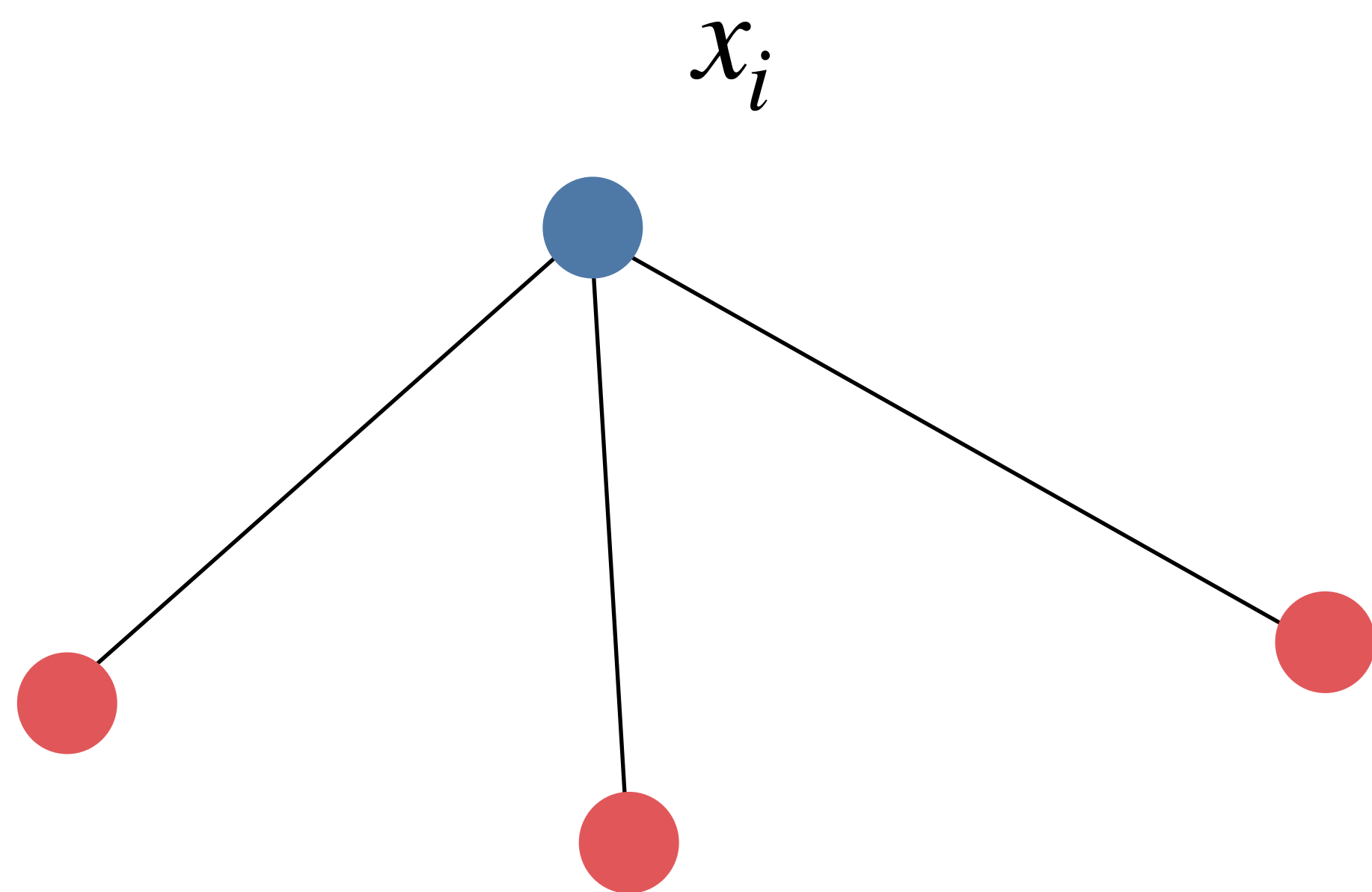
Create Message



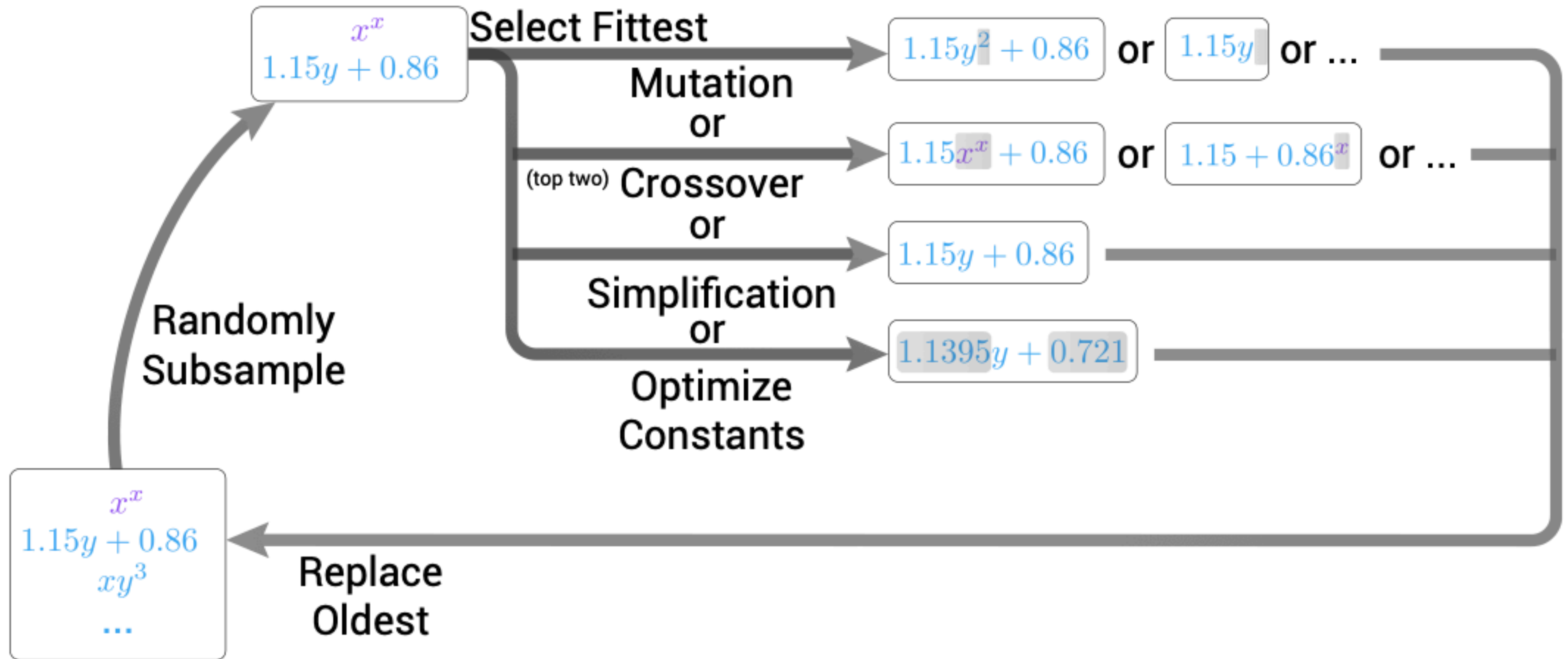
Aggregate



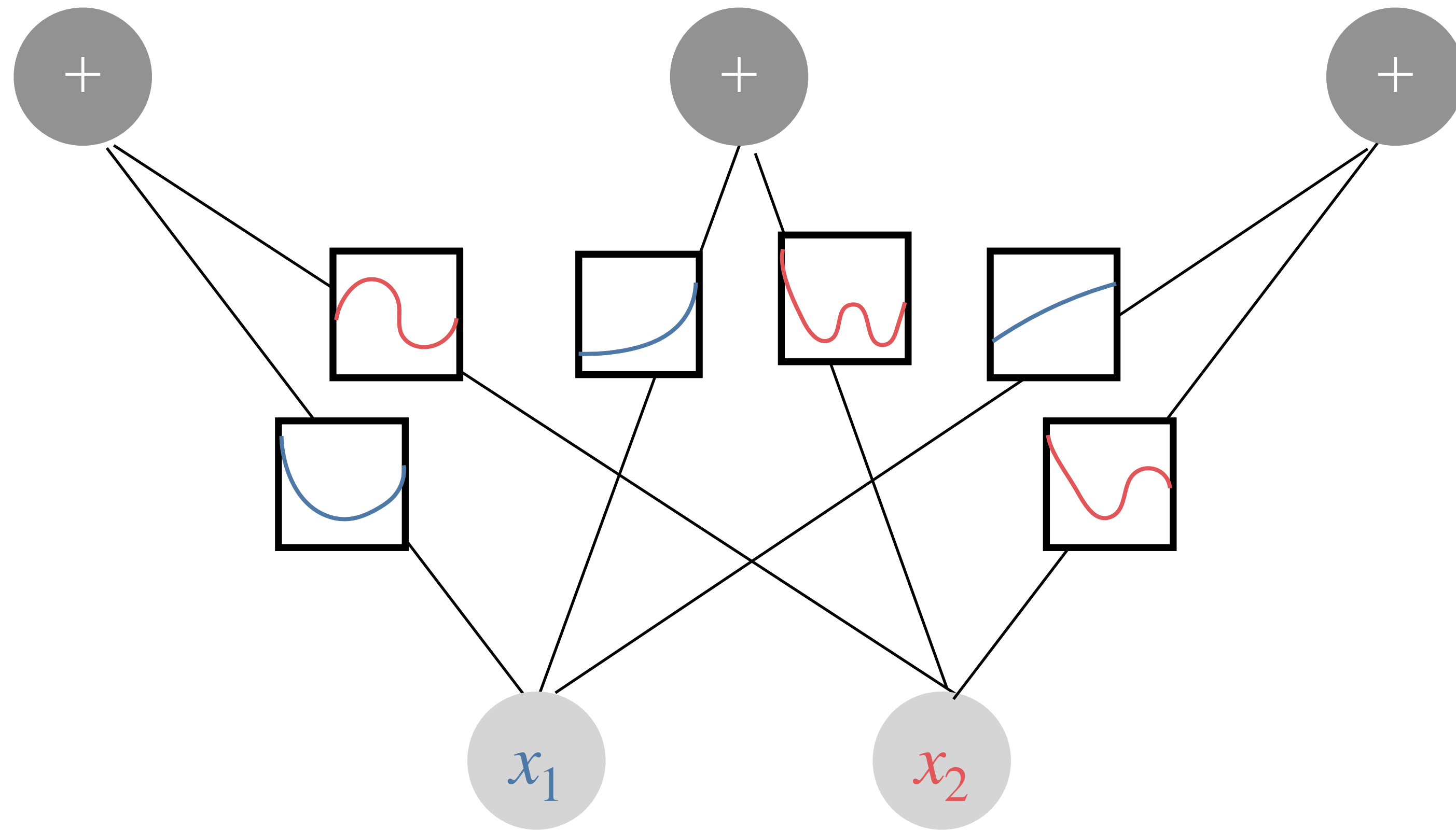
Update



Symbolic Regression



Kolmogorov Arnold Networks

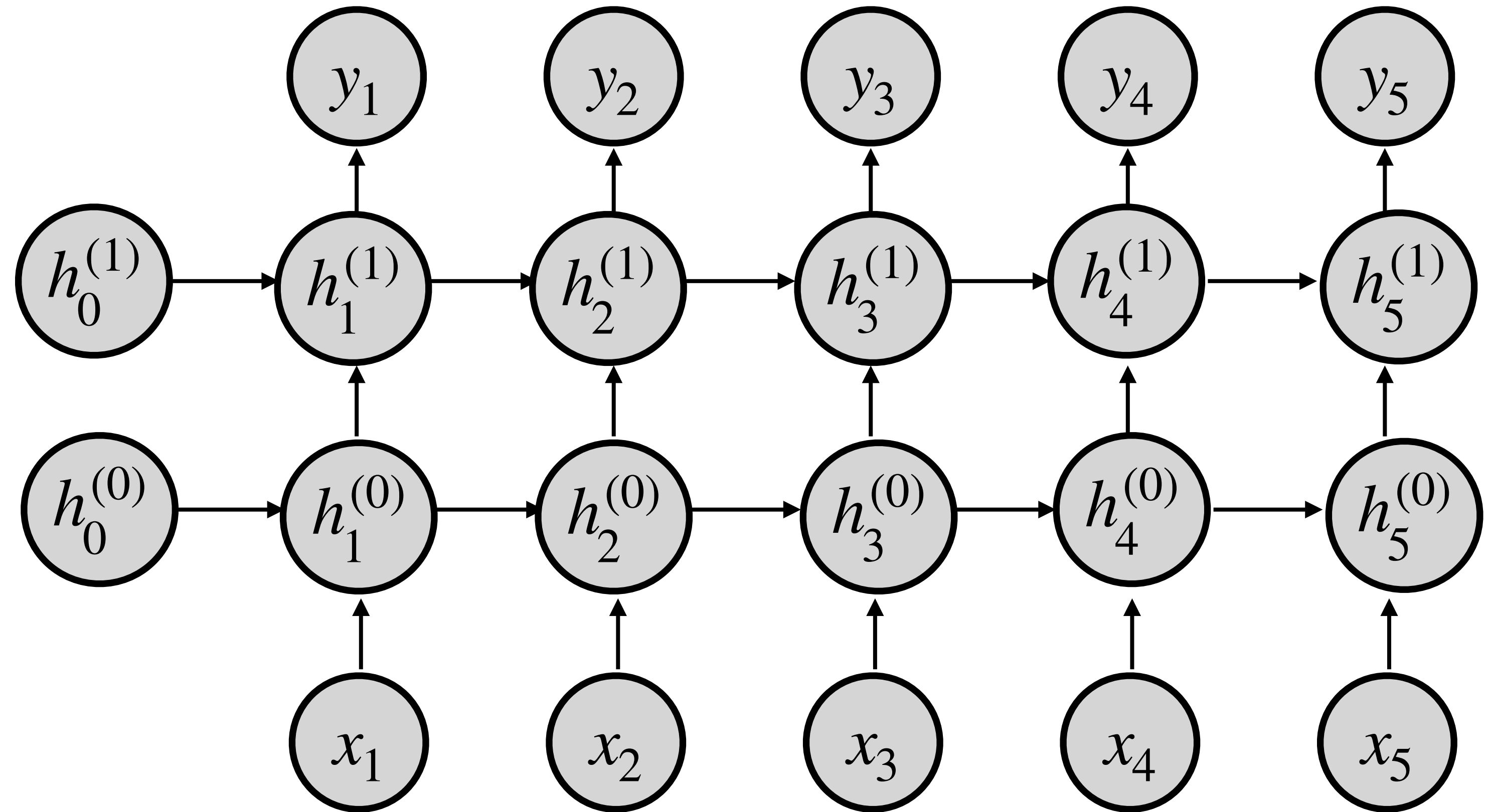


Recurrent Neural Nets

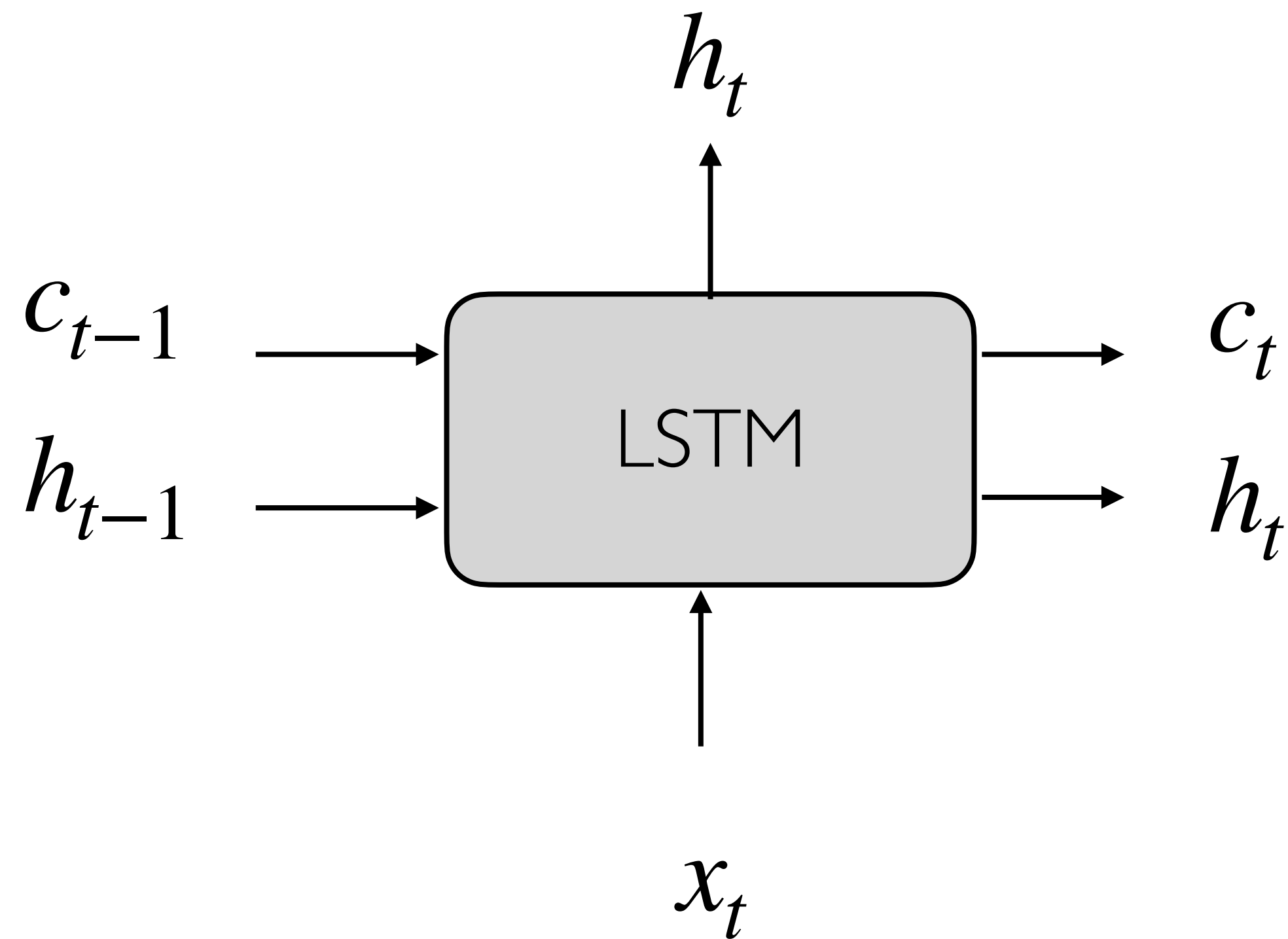
outputs:

hidden:

inputs:



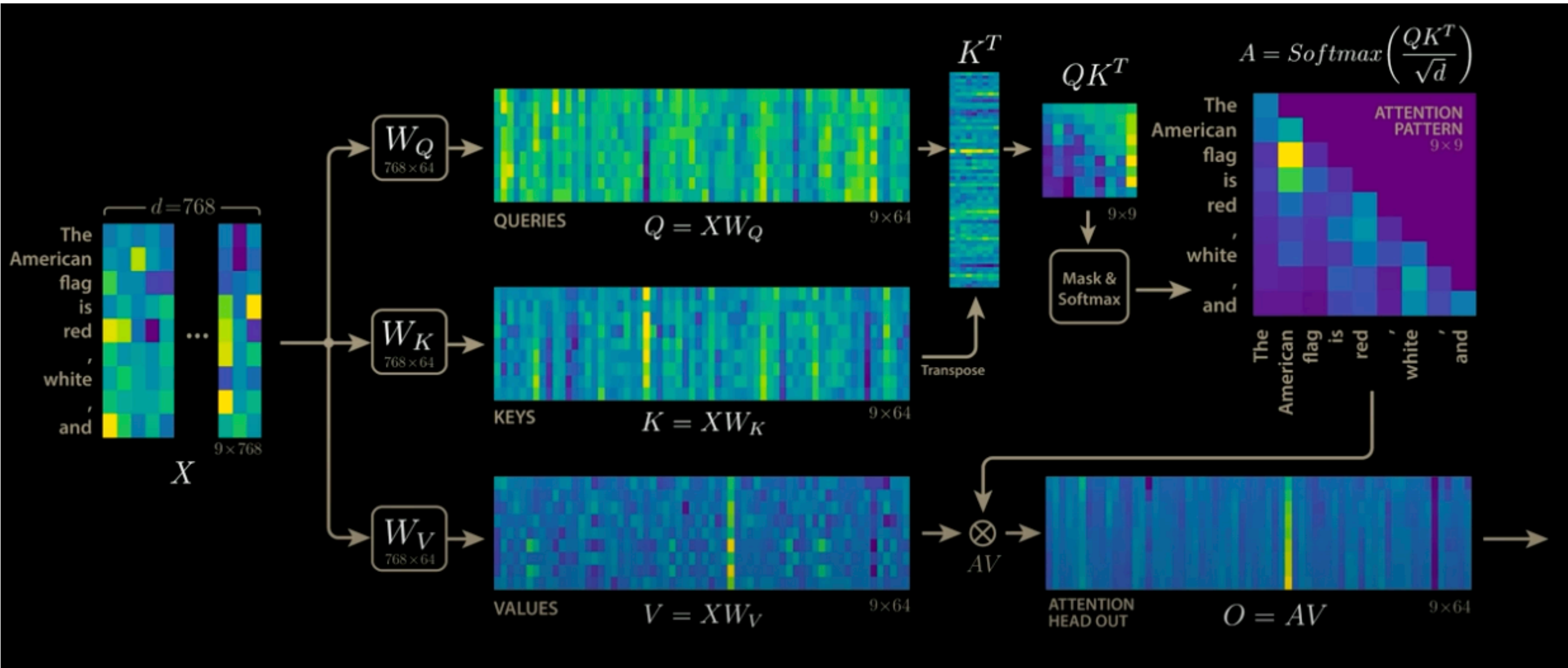
GRU and LSTM



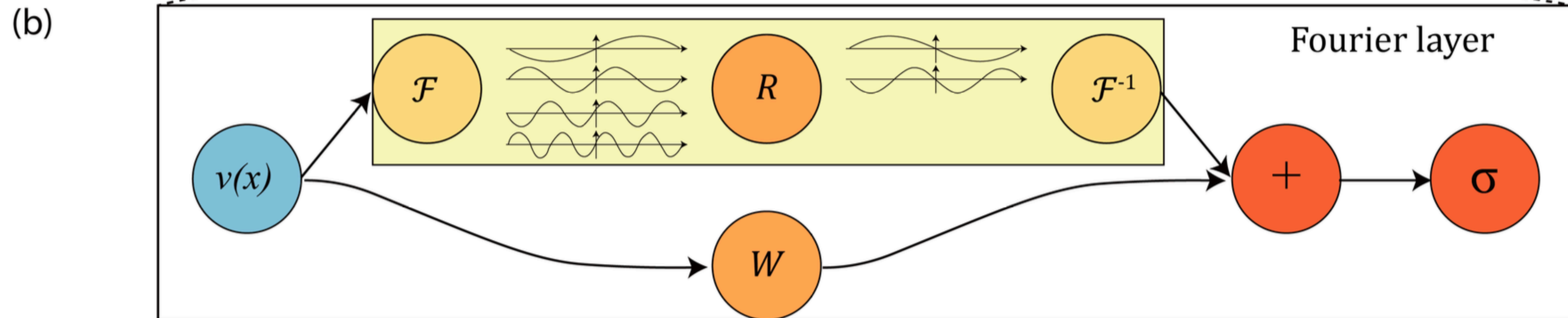
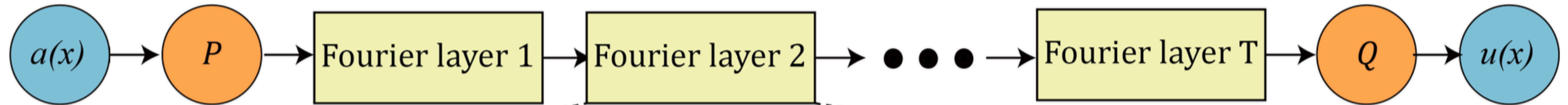
sigmoid function

$$\Gamma = \sigma(W_x x^{(t)} + W_h h^{(t-1)} + b)$$

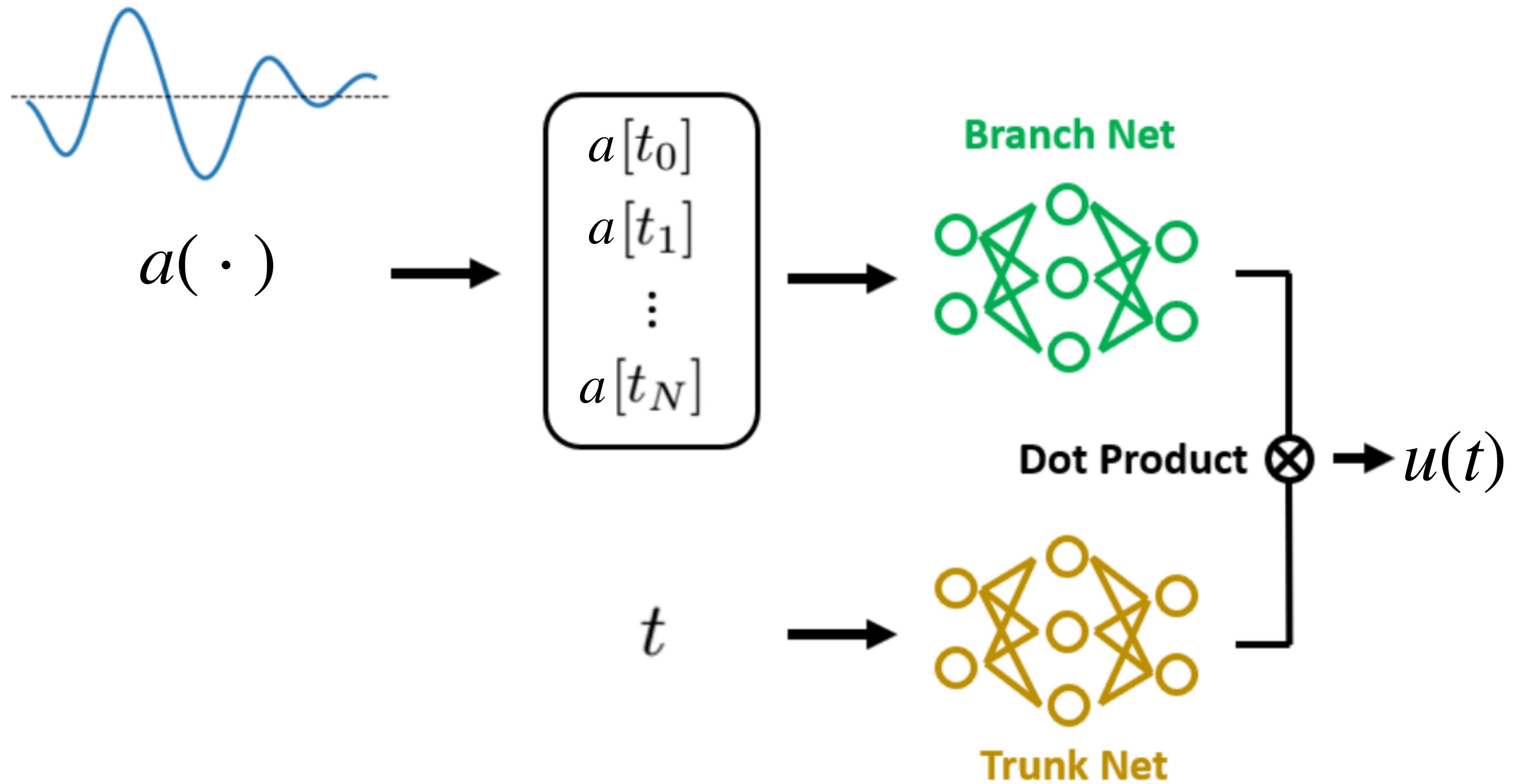
Transformers



Fourier Neural Operators



DeepONet



Parting thought