

Course #3:

**Auto-encoders and
Recurrent Neural Networks**

Roadmap

- Recap from course #2
- Auto-encoders
- Recurrent Neural Networks

Lecture. #2

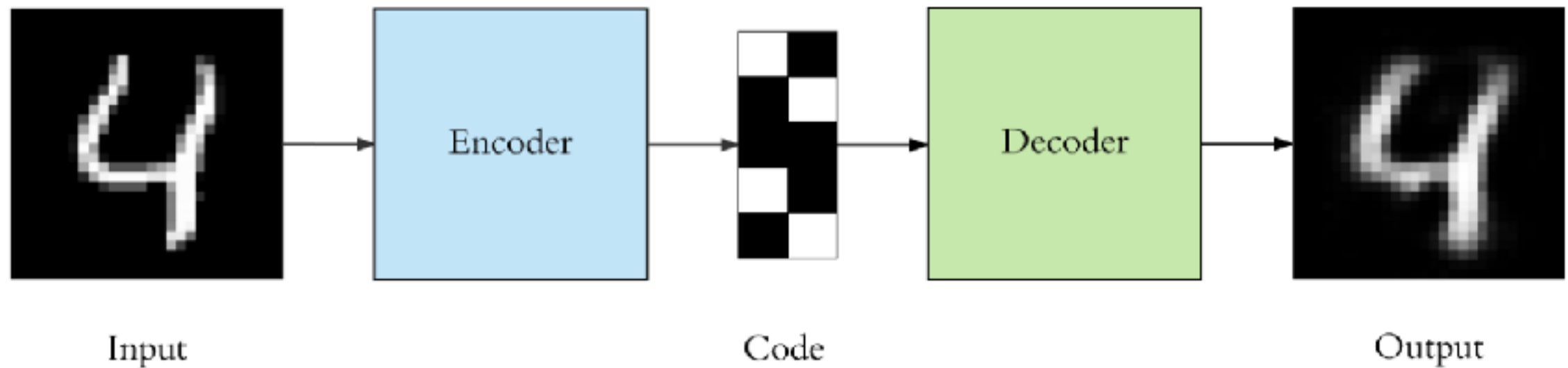
Things to know

- Convolution layers
- Pooling layers
- Activation layers
- Dropout layers
- Padding and stride
- Fine-tuning
- Over-fitting
- Data augmentation
-

PyTorch Lightning

Auto-encoders

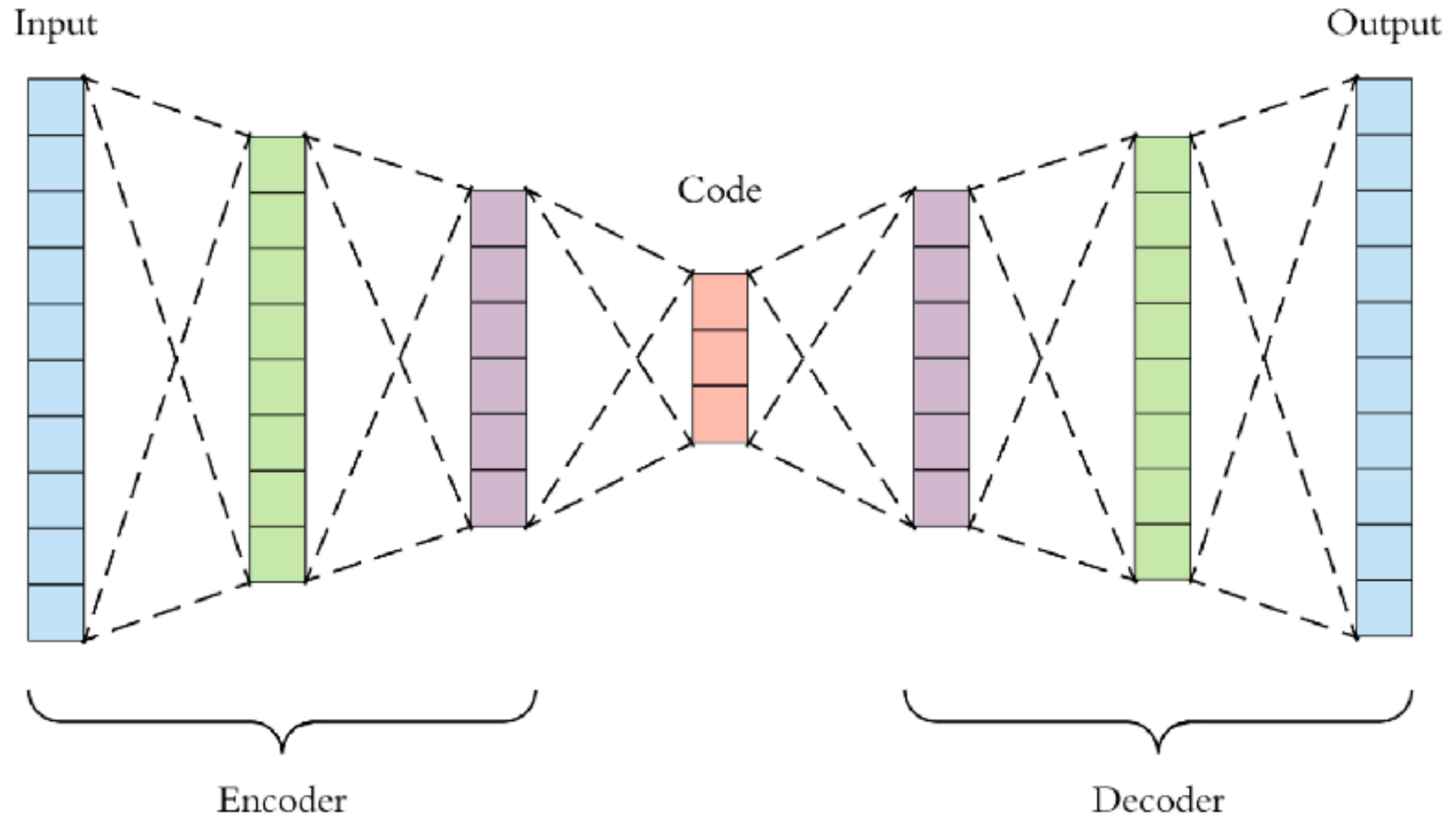
Auto-encoders



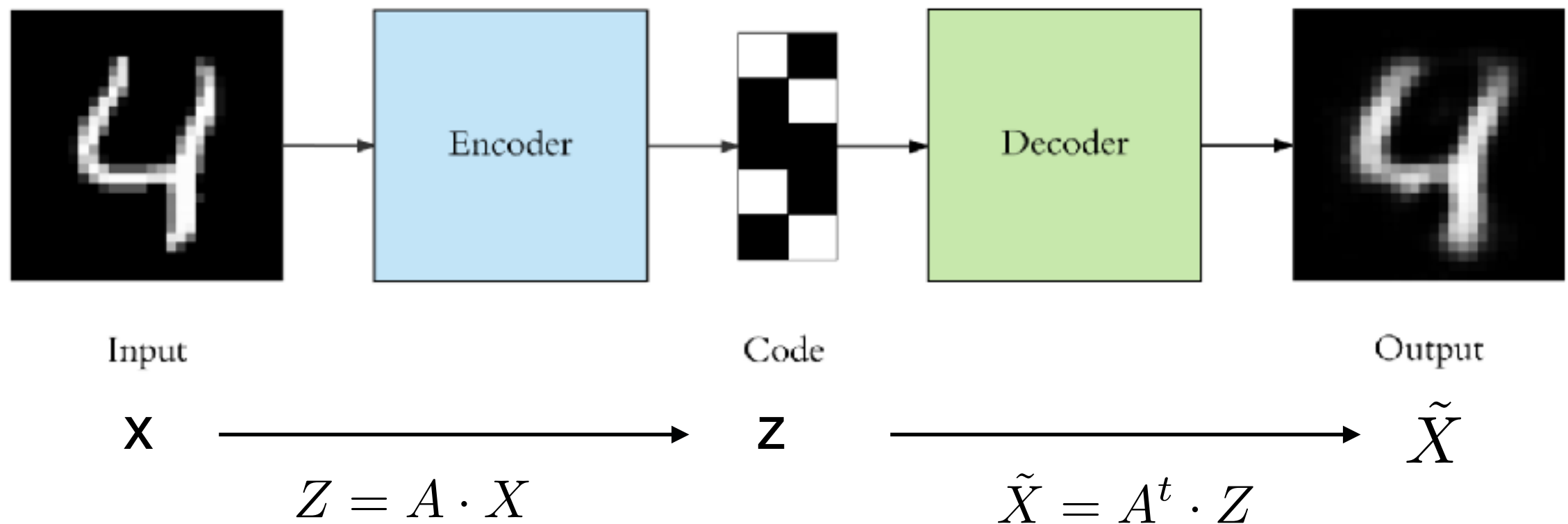
Output with the same shape as the input

Application ?

Dense auto-encoders



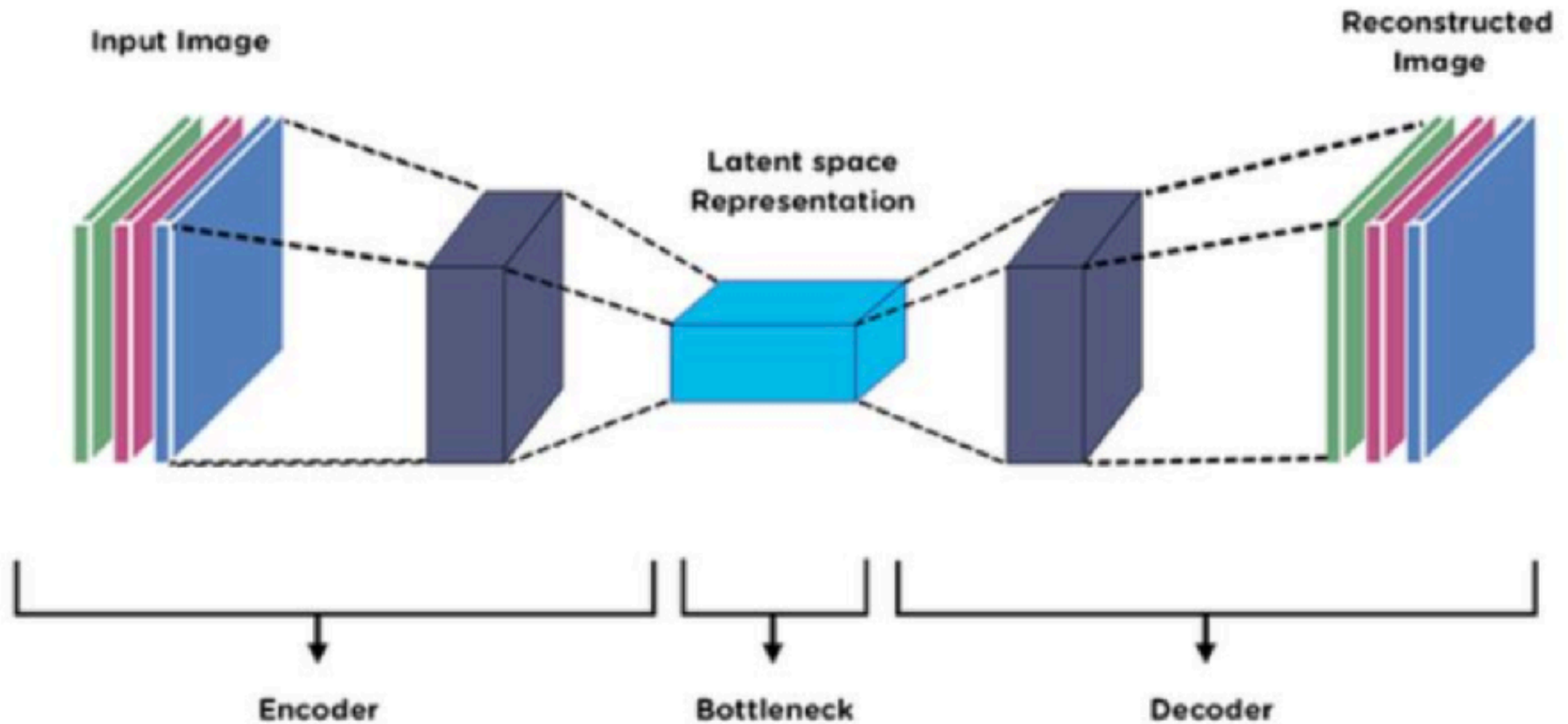
PCA/EOF



PCA as a linear auto-encoder architecture.

Which additional constraint ?

Convolutional auto-encoders

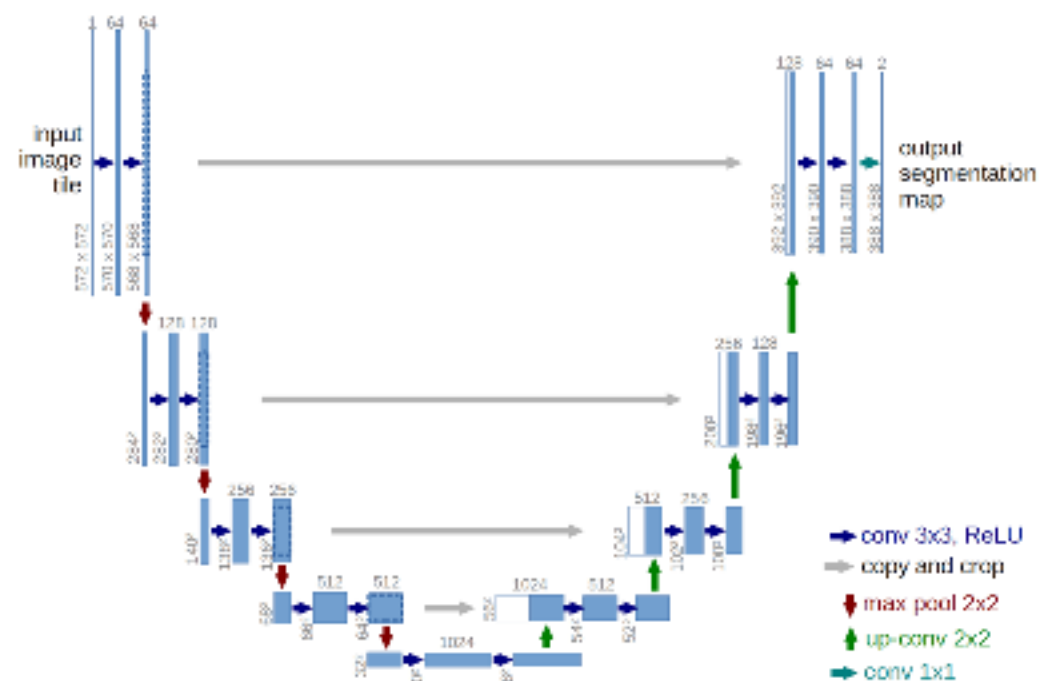


Convolutional AE Zoo

Many applications do not require a low-dimensional representation
(e.g., denoising, interpolation, super-resolution,...)



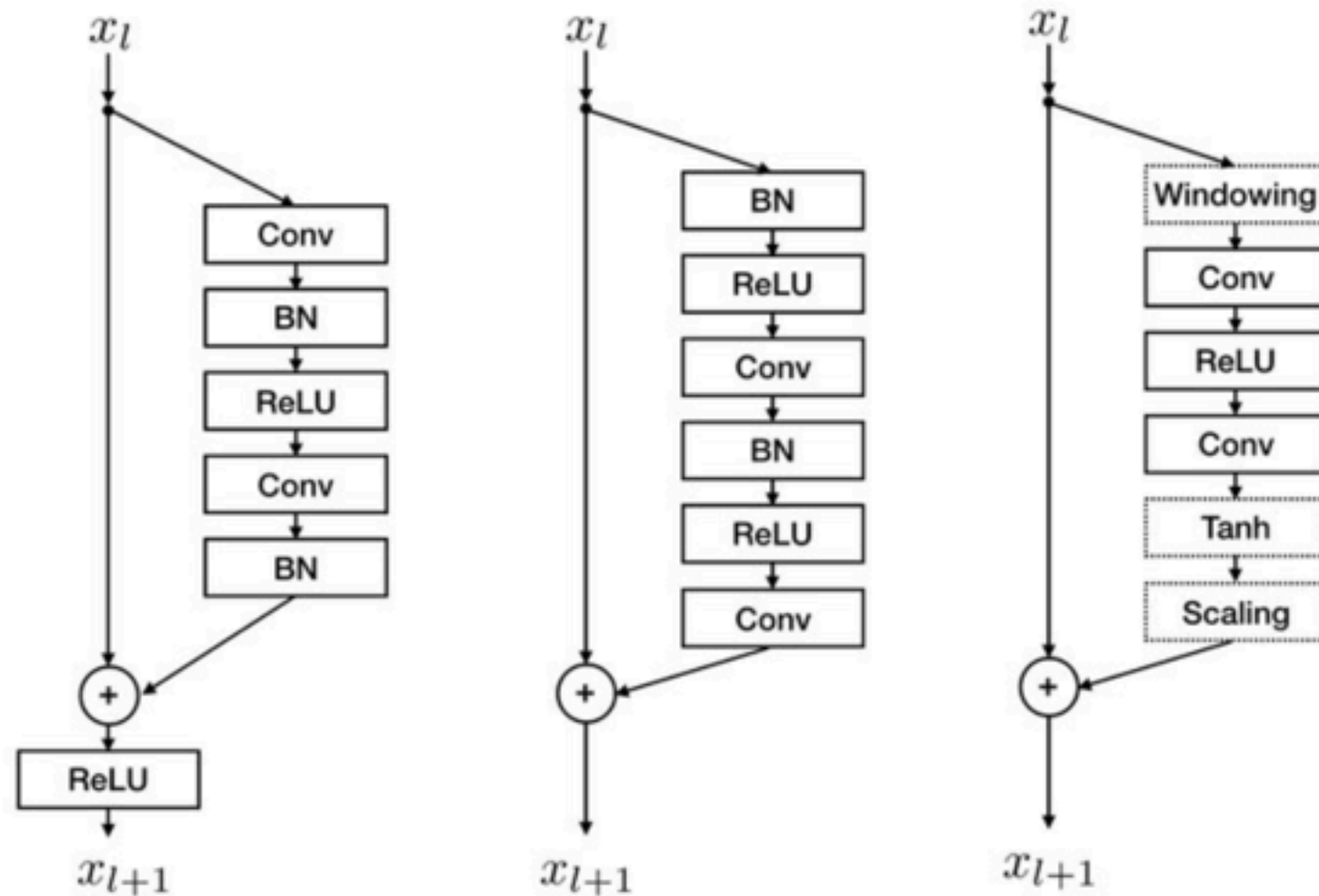
<https://arxiv.org/pdf/1606.08921.pdf>



UNet

Ronneberger
et al., 2015

Convolutional AE Zoo



Residual Block
(ResNet)

Rousseau et al.,
2019

Often used to address vanishing gradients (“very” deep networks)

Auto-encoders for image denoising and image generation

https://github.com/CIA-Oceanix/DLCourse_MOi_2022/blob/main/notebooks/notebook_PytorchLightning_MNIST_AutoEncoder_students.ipynb

Question 2. Fill in the architecture of the convolutional block of the encoder module to train a convolutional auto-encoder

Question 2. Modify the Conv. AE code to address image denoising and/or image interpolation

Question 3. Modify the code to test a linear auto-encoder (cf. AE and PCA)

ConvAE architectures for Ocean Dynamics ?

ConvAE & Ocean Dynamics

JAMES | Journal of Advances in Modeling Earth Systems

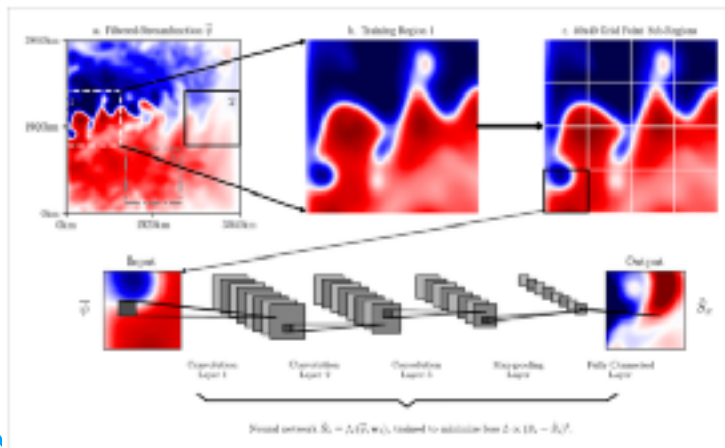
RESEARCH ARTICLE
10.1029/2018MS001472

Key Points:
• We successfully use convolutional neural networks to predict unresolved turbulent processes and subsurface velocities.
• The neural networks generalize to

Applications of Deep Learning to Ocean Data Inference and Subgrid Parameterization

Thomas Bolton¹ and Laure Zanna¹

¹Department of Physics, University of Oxford, Oxford, UK



[https://
agupubs.onlinelibrary.wiley.co
m/doi/epdf/
10.1029/2018MS001472](https://agupubs.onlinelibrary.wiley.com/doi/epdf/10.1029/2018MS001472)

JAMES | Journal of Advances in Modeling Earth Systems[®]

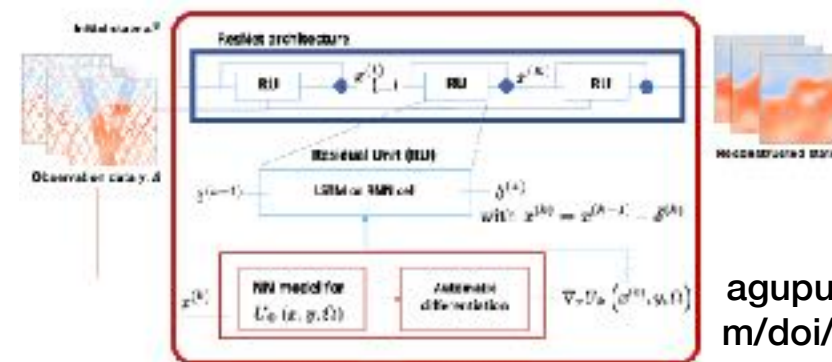
RESEARCH ARTICLE
10.1029/2021MS002572

Key Points:
• We develop a end-to-end neural architecture to design variational data assimilation models and solvers.
• We may jointly calibrate all the trainable components of the proposed scheme to optimize a data

Learning Variational Data Assimilation Models and Solvers

R. Fablet¹, B. Chapron², L. Drumetz¹, E. Mémin³, O. Panneton⁴, and F. Rousseau⁵

¹IMT Atlantique, UMR CNRS Lab-STICC, Brest, France, ²IFREMER, UMR CNRS LOOS, Brest, France, ³INRIA Rennes, UMR CNRS IRISA, Rennes, France, ⁴INRIA-ENM, UMR CNRS CNRM, CERFACS, Toulouse, France, ⁵IMT Atlantique, UMR INSERM Latim, Brest, France



[https://
agupubs.onlinelibrary.wiley.co
m/doi/10.1029/2021MS002572](https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2021MS002572)

JAMES | Journal of Advances in Modeling Earth Systems

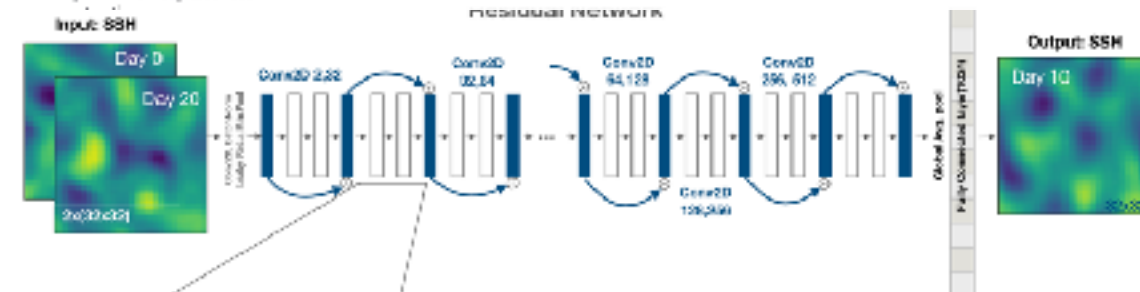
RESEARCH ARTICLE
10.1029/2019MS001965

Key Points:
• The efficacy of Deep Learning in exploiting sparse sea surface height (SSH) data is demonstrated in a quasigeostrophic model.
• Residual Neural Networks are superior to linear and dynamical interpolation techniques for SSH

A Deep Learning Approach to Spatiotemporal Sea Surface Height Interpolation and Estimation of Deep Currents in Geostrophic Ocean Turbulence

Georgy E. Manucharyan¹, Lia Siegelman², and Patrice Klein^{3,4}

¹School of Oceanography, University of Washington, Seattle, WA, USA, ²The Population Laboratory, California Institute of Technology, Pasadena, CA, USA, ³Laboratoire de Mécanisme Dynamique, Ecole Normale Supérieure, CNRS, Paris, France, ⁴Laboratoire d'Océanographie Physique et Spatiale, IFREMER, CNRS, Brest, France



[https://agupubs.onlinelibrary.wiley.com/
doi/epdf/10.1029/2019MS001965](https://agupubs.onlinelibrary.wiley.com/doi/epdf/10.1029/2019MS001965)

DINCAE 2.0: multivariate convolutional neural network with error estimates to reconstruct sea surface temperature satellite and altimetry observations

Alexander Barth¹, Aida Alvera-Azcárate¹, Charles Troupin¹, and Jean-Marie Beckers¹

GHEF, University of Liège, Liège, Belgium

Received: 18 Oct 2021 – Discussion started: 15 Nov 2021 – Revised: 10 Feb 2022 – Accepted: 17 Feb 2022 – Published: 15 Mar 2022

<https://gmd.copernicus.org/articles/15/2183/2022/>

ConvAE & Ocean Dynamics

Literature review

Considered papers:

Topic#1 <https://gmd.copernicus.org/articles/15/2183/2022/>

Topic#2 <https://agupubs.onlinelibrary.wiley.com/doi/epdf/10.1029/2018MS001472>

Topic#3 <https://agupubs.onlinelibrary.wiley.com/doi/epdf/10.1029/2019MS001965>

Topic#4 <https://arxiv.org/abs/2010.04663>

Questions:

- Which problem ?
- Which convolutional architecture ?
- Comments ?

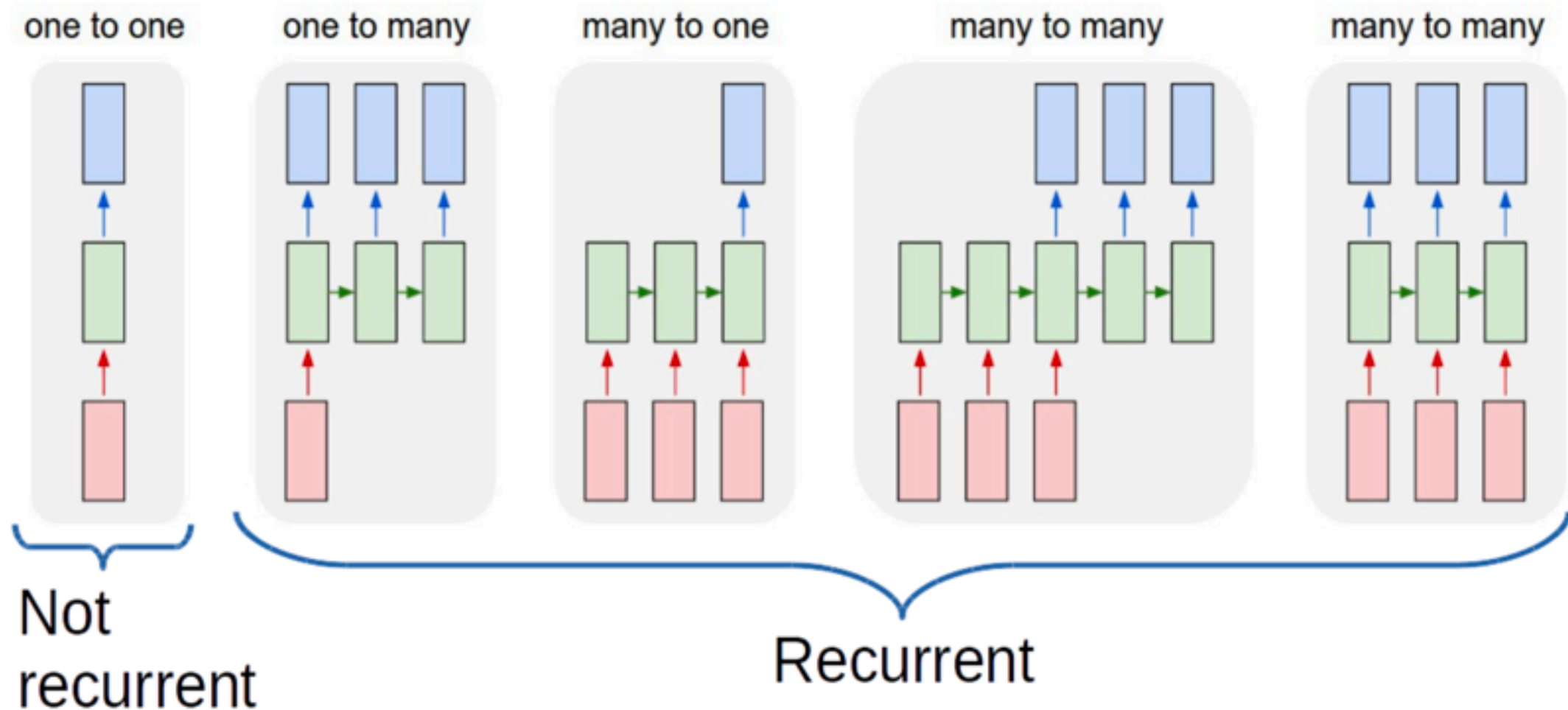
Lecture. #3

Things to know (AE)

- Auto-encoder
- Latent variable
- UNet
- ResNet

Recurrent Neural Networks

Recurrent Neural Networks



Applications:

- Time-series forecasting
- Audio processing
- Translation
-

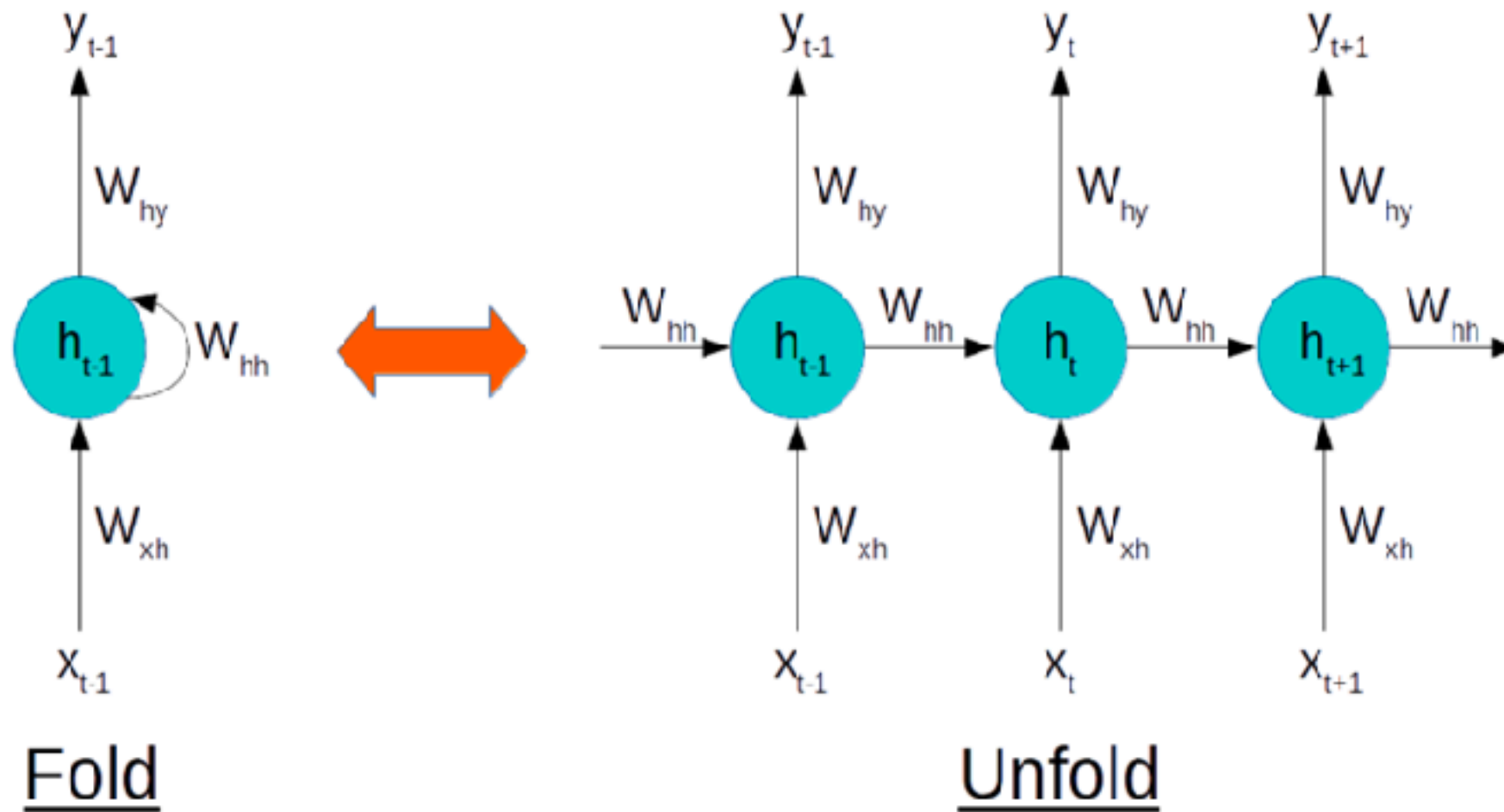
RNN: Underlying formulation (similar to state-space representation)

$x \rightarrow$ input vector

$y \rightarrow$ output vector

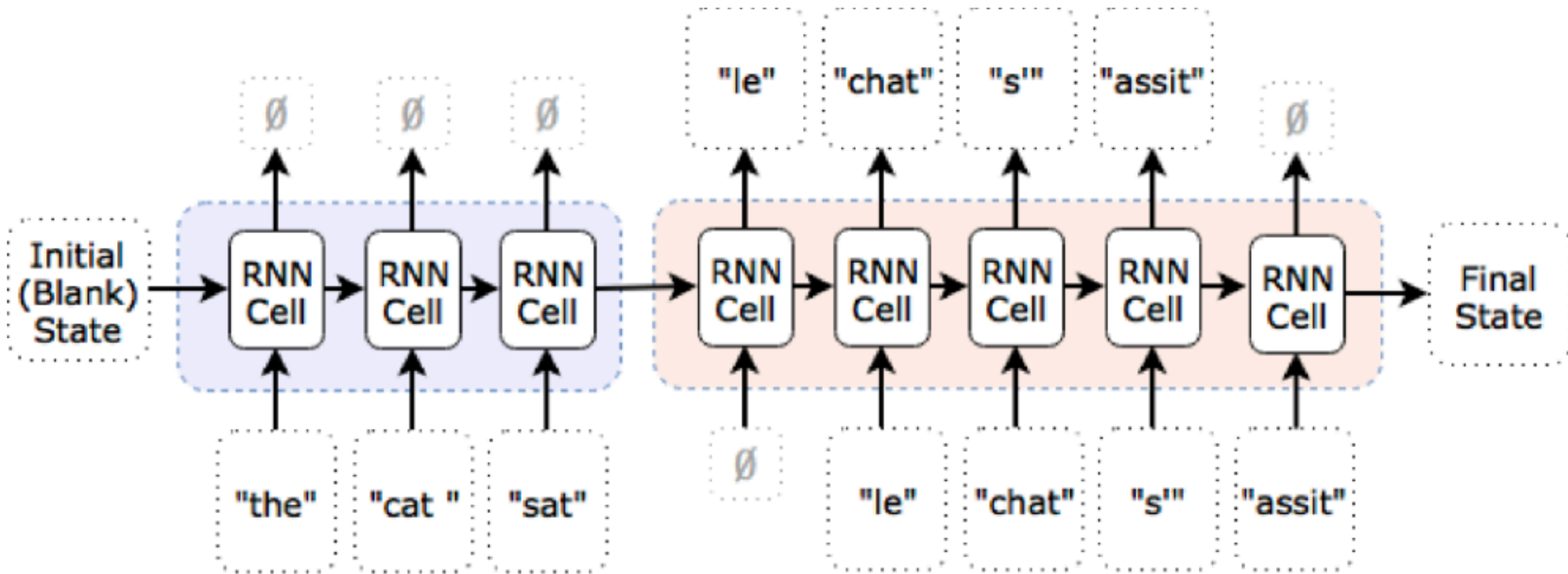
$h \rightarrow$ hidden state

$t \rightarrow$ index (time, index, etc...)



$$h_t = f(W_{hh}h_{t-1} + W_{xh}x_t) \text{ and } y_t = g(W_{hy}h_t)$$

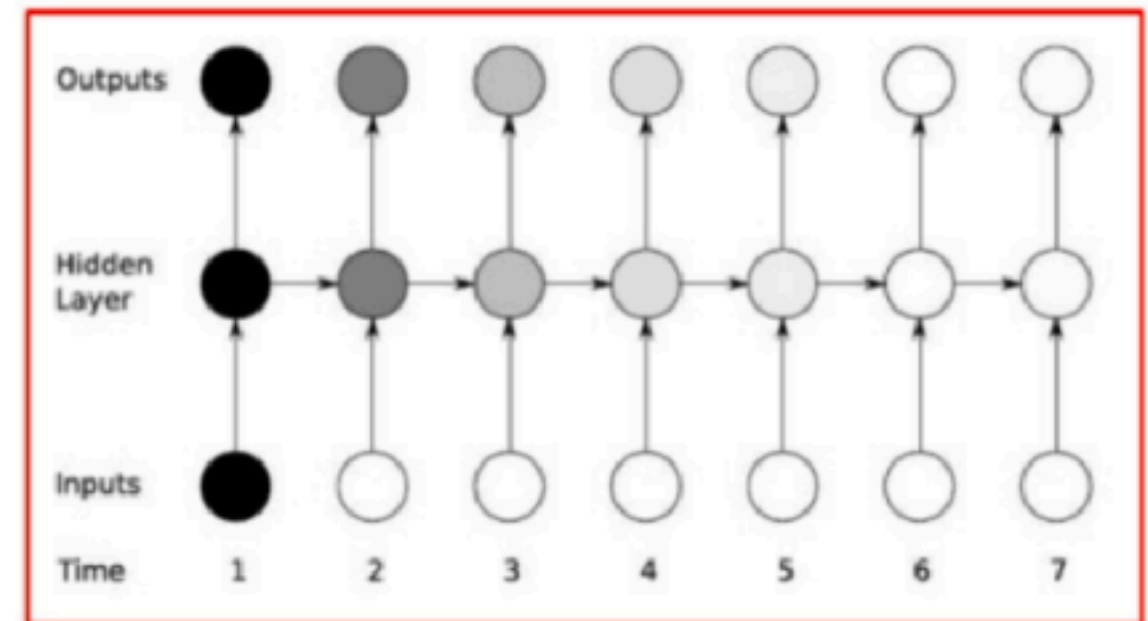
RNN: Underlying formulation (similar to state-space representation)



Classic RNN Architectures

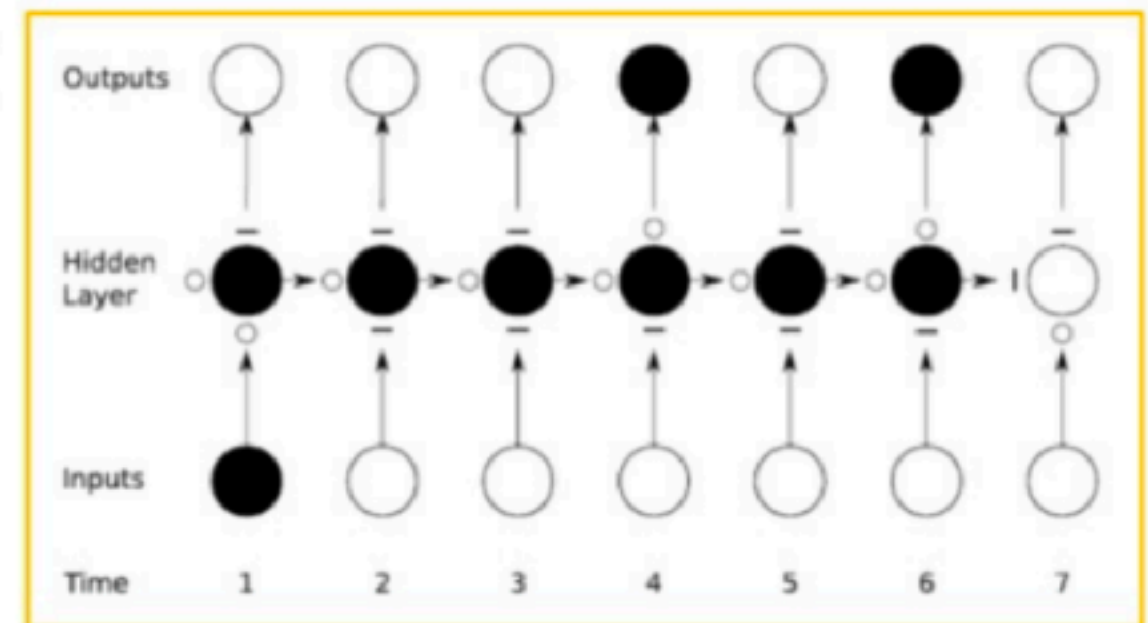
- **Conventional RNN with sigmoid**

- The sensitivity of the input values decays over time
- The network forgets the previous input



- **Long-Short Term Memory (LSTM)** [2]

- The cell remember the input as long as it wants
- The output can be used anytime it wants



Dense and convolutional versions of LSTM and GRU exist
(depending on the structure of the hidden state)

Short-term forecasting application (L63 case-study)

[https://github.com/CIA-Oceanix/DLCourse_MOi_2022/blob/main/notebooks/
notebook_Forecasting_L63_students.ipynb](https://github.com/CIA-Oceanix/DLCourse_MOi_2022/blob/main/notebooks/notebook_Forecasting_L63_students.ipynb)

Lecture. #3

Things to know (RNN)

- Recurrent Neural Network
- LSTM
- Unfloded and folded representations