

Chapter ML:IX (continued)

IX. Deep Learning

- ❑ Elements of Deep Learning
- ❑ Convolutional Neural Networks
- ❑ Autoencoder Networks
- ❑ Recurrent Neural Networks
- ❑ Long-Term Dependencies
- ❑ RNNs for Machine Translation
- ❑ Attention Mechanism
- ❑ Self Attention and Transformers
- ❑ Transformer Language Models

Long-Term Dependencies

Vanishing Gradient Problem

Long-Term Dependencies

RNN with Long Short-Term Memory (LSTM)

[*SKIPPED*]

Remarks:

- ❑ LSTM is a recurrent neural network architecture that is very efficient at remembering long term dependencies and that is less vulnerable to the vanishing gradient problem.

Long-Term Dependencies

RNN with Gated Recurrent Units (GRU)

[*SKIPPED*]