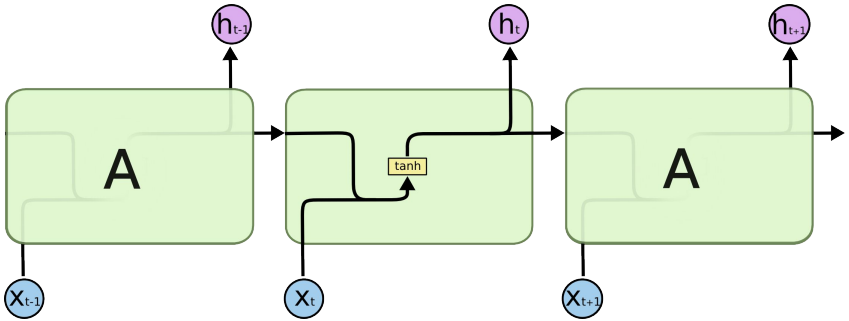
DAY 5

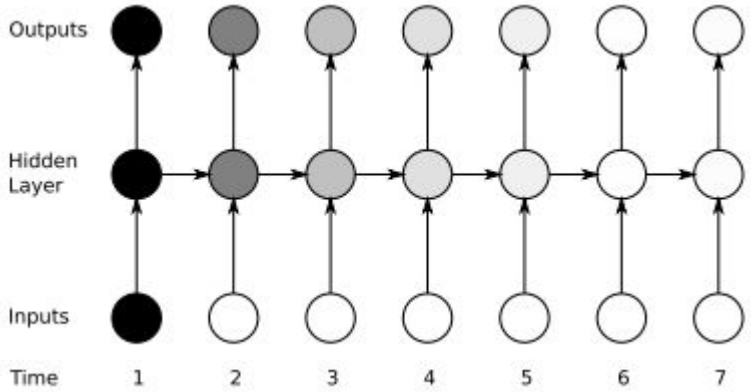
RNN

RNN

- Basic RNNs suffer from:
  - Vanishing or exploding gradients
  - Hard to remember connections with long sequences



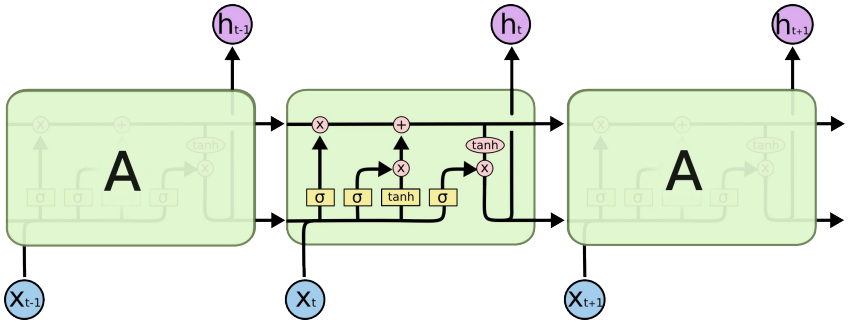
Source: Christopher Olah,Understanding LSTM Networks, 2015 (Blog)



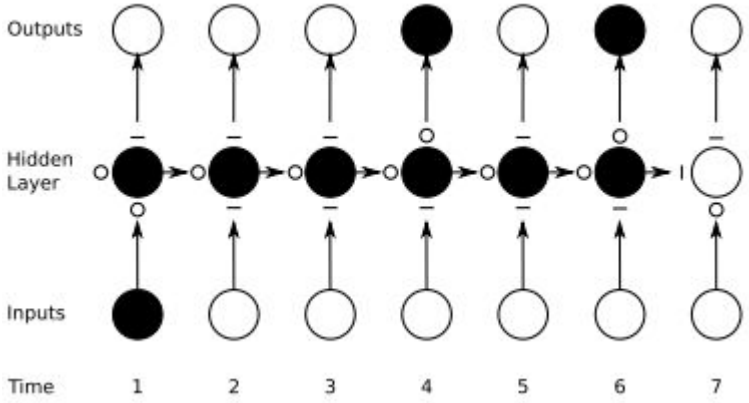
Source: Alex Graves,
Supervised Sequence Labelling with Recurrent Neural Networks, 2012 (Preprint)

RNN

- Long Short Term Memory (LSTM)
  - Solves some of the shortcomings of RNNs
  - Better learning of complex sequences



Source: Christopher Olah,Understanding LSTM Networks, 2015 (Blog)



Source: Alex Graves,
Supervised Sequence Labelling with Recurrent Neural Networks, 2012 (Preprint)

RNN

- Cell types (Python API)
  - **BasicRNNCell** → tf.nn.rnn_cell.BasicRNNCell
    - `activation=tanh`
    - `num_units`
  - **BasicLSTMCell** → tf.nn.rnn_cell.BasicLSTMCell
    - `forget_bias`
    - Zaremba et al. (2014)
  - **GRUCell** (Gated Recurrent Unit) → tf.nn.rnn_cell.GRUCell
    - Cho et al. (2014)
  - **LSTMCell** → tf.nn.rnn_cell.LSTMCell
    - Default Hochreiter and Schmidhuber (1997) implementation (use_peepholes=False)
    - Sak et al. (2014) implementation (use_peepholes=True)

Source: Taegyun Jeon. Electricity price forecasting with Recurrent Neural Networks (SlideShare)

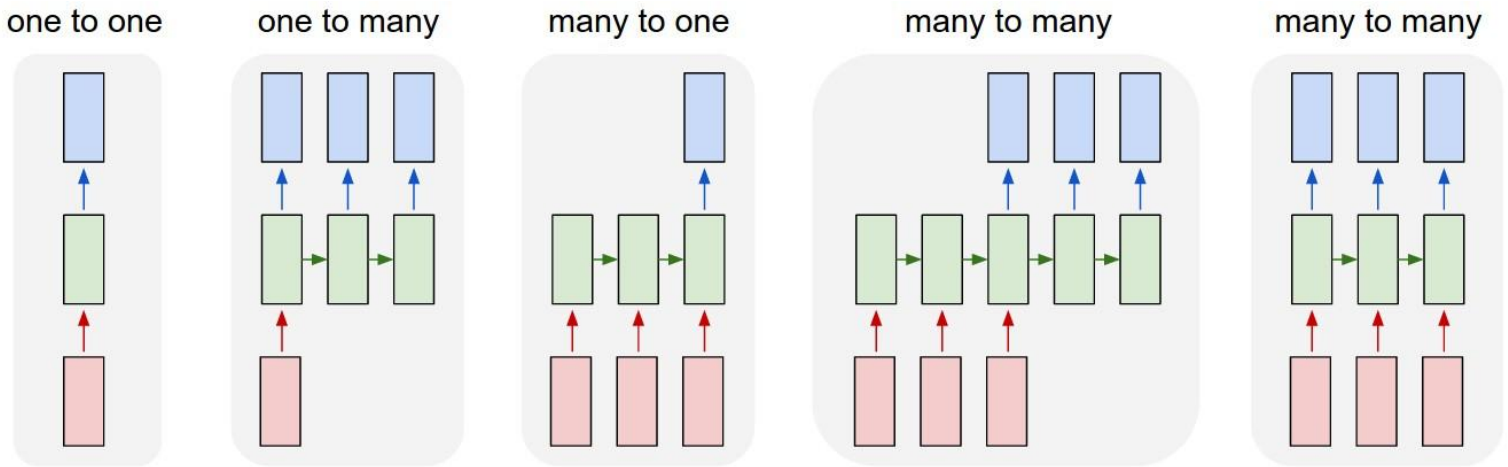
RNN

- Cell wrappers (Python API)
  - **MultiRNNCell** → tf.nn.rnn_cell.MultiRNNCell
    - RNN cell composed sequentially of a number of RNNCells
  - **Dropout** → tf.nn.rnn_cell.DropoutWrapper
    - Dropout probability between layers
    - `input_keep_prob=1.0, output_keep_prob=1.0`
  - **Embedding** → tf.nn.rnn_cell.EmbeddingWrapper
    - Operator adding input embedding to the given cell
    - `embedding_classes, embedding_size`

Source: Taegyun Jeon. Electricity price forecasting with Recurrent Neural Networks (SlideShare)

RNN

- Some applications



Source: Andrej Karpathy. The Unreasonable Effectiveness of Recurrent Neural Networks (Blog)

RNN

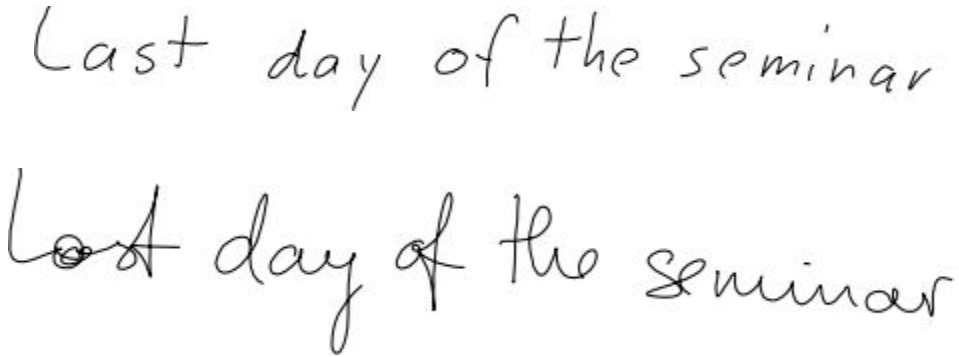
- TensorFlow tutorials
  - Vector Representation of Words (Link)
    - In depth explanation of the word2vec model
    - Expands the explanation in yesterday's Language lecture
  - Recurrent Neural Networks (Link)
    - Firsts steps at training a RNN for text
    - The model predict the next word given a sentence
    - Stack of multiple BasicLSTMCell
    - Different model complexities and parameters to play with

RNN

- TensorFlow tutorials
  - Sequence-to-Sequence Models (Link)
    - Training a RNN for English to French translation
    - Sequence to sequence implementation
  - SyntaxNet (Link, Google Research Blog)
    - Foundation for Natural Language Understanding (NLU) systems
    - Tags each word with a part-of-speech (POS) tag that describes the word's syntactic function

RNN

- RNN handwriting generation demo (A. Graves, 2013)
  - Online demo



  - TensorFlow implementation
    - Code, Blog

RNN

- Text generation (A. Karpathy, 2015)
  - Code, torch-rnn
  - TensorFlow implementation
    - Code
- Music generation (Magenta)
  - Application of language modeling to melody generation
  - Code
- More ‘common’ tasks such as regression and classification
  - Interesting for time series data
  - Sentiment analysis

TensorFlow resources

- Main documentation
  - Very well written
  - Constantly including new tutorials and How-Tos
- Some resources
  - Collection of tutorials, papers, projects, resources, books, etc. → Awsome TensorFlow
  - Pretrained models for MatConvNet
    - Can be loaded to TF using scipy.io
  - Deep Reinforcement Learning → Code
  - ResNets → Code
  - Deep Learning Algorithms → Code