Today: Outline

- **Pre-lecture Material**
- **Recurrent Neural Networks**

- Reminders:

  Problem Set 1, due: Oct 12 by midnight
  Midterm Exam, in class, Oct 20 (Practice problems will be posted)

- Announcement:

  No class on Oct 13 per BU Calendar
  (Substitute Mon Schedule of Classes)

# Neural Networks V

Pre-lecture Material

AI Generated Trailer

How did Watson generate a movie trailer?

- ☐ Using human expertise only
- ☐ Using tools humans use to generate trailers
- ☐ By identifying salient regions in the movie
- ☐ By detecting times where actions and emotions are predicted

AI Generated Trailer

- Analyze a movie and generate a trailer automatically

- How?
  Detecting salient moments
  e.g. action/emotions

MORGAN Full Movie - Salient Sentiment Change Detection
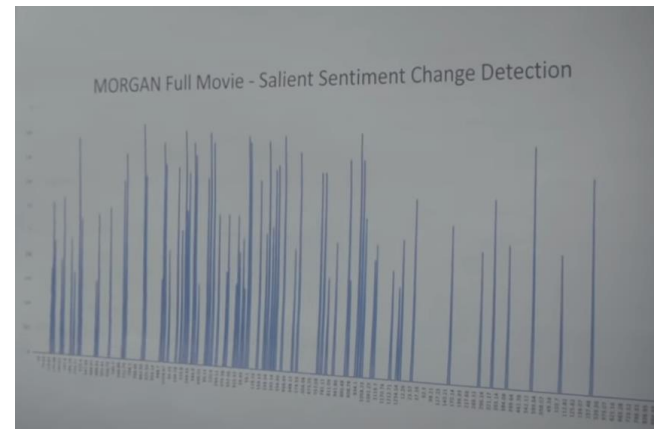
Detecting Salient Regions

- Two sample actions:
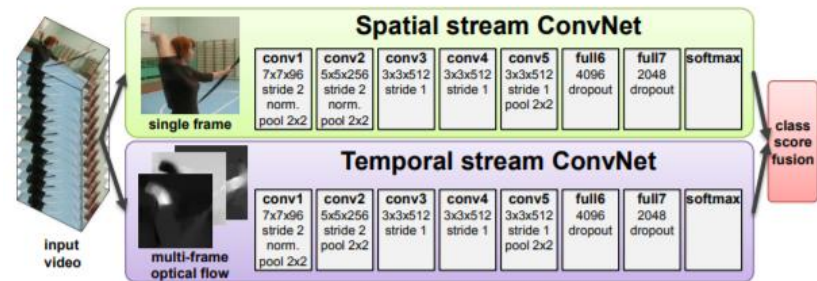
Handstand Walking   *Ice Dancing*

AI Generated Match Highlights

- IBM produces the official match highlights of Wimbledon and US Open tennis tournaments.

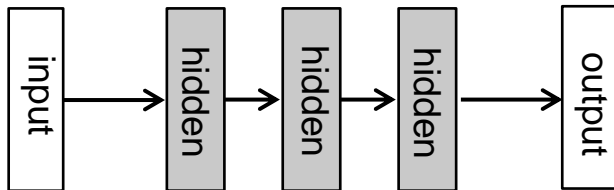- https://www.usopen.org/en_US/video/2017-08-31/1504233424.html

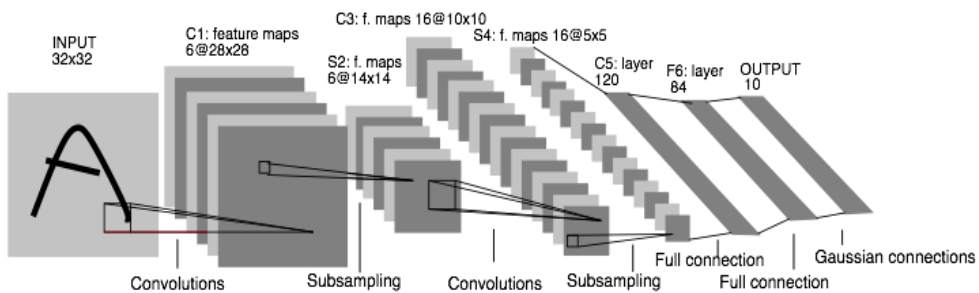- Multi-modal System

- Bias Considerations
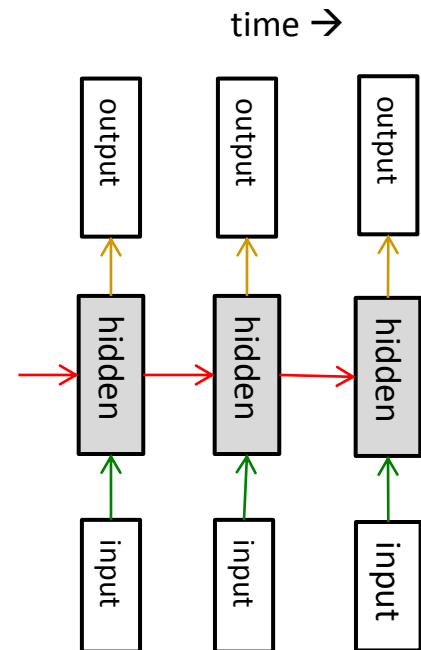
Network architectures

Feed-forward

Fully connected

Convolutional

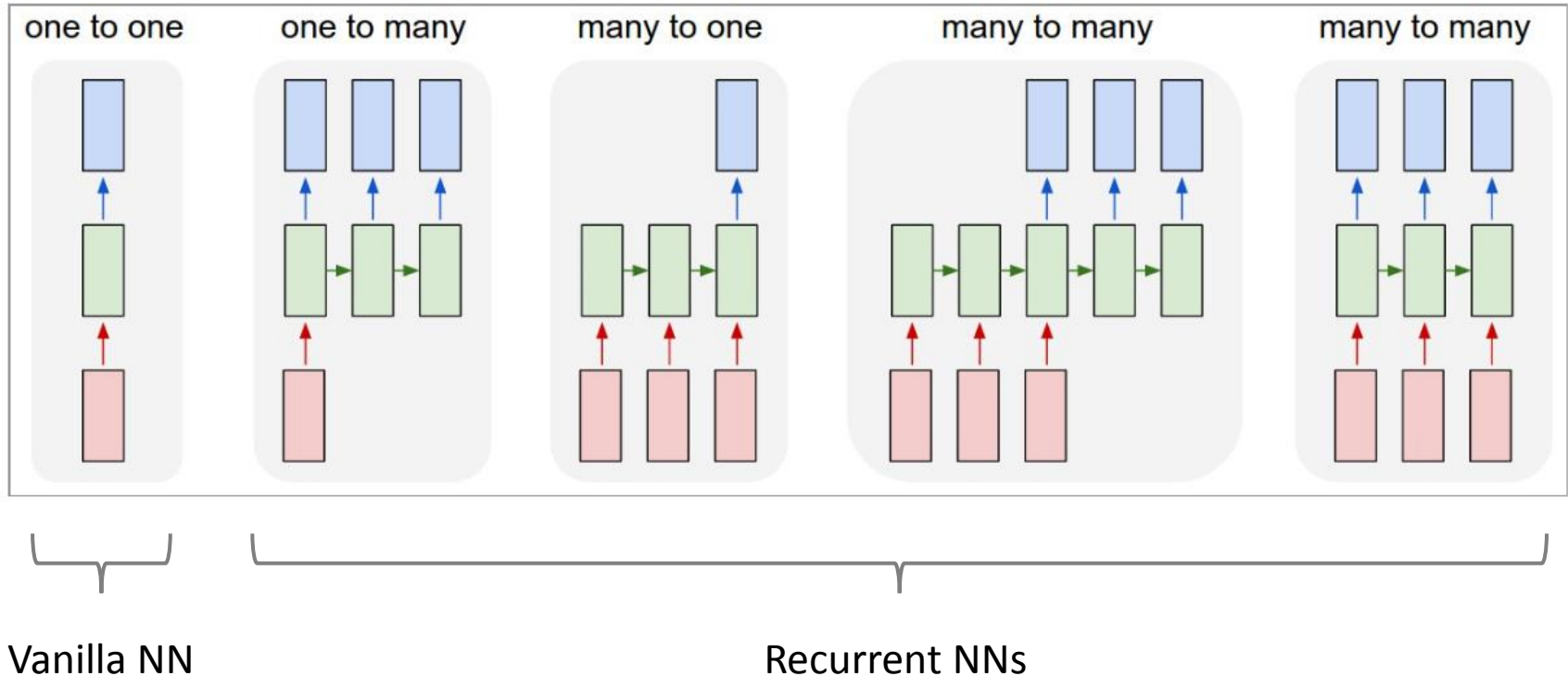
Recurrent

time →

Neural Networks V

Recurrent Neural Networks

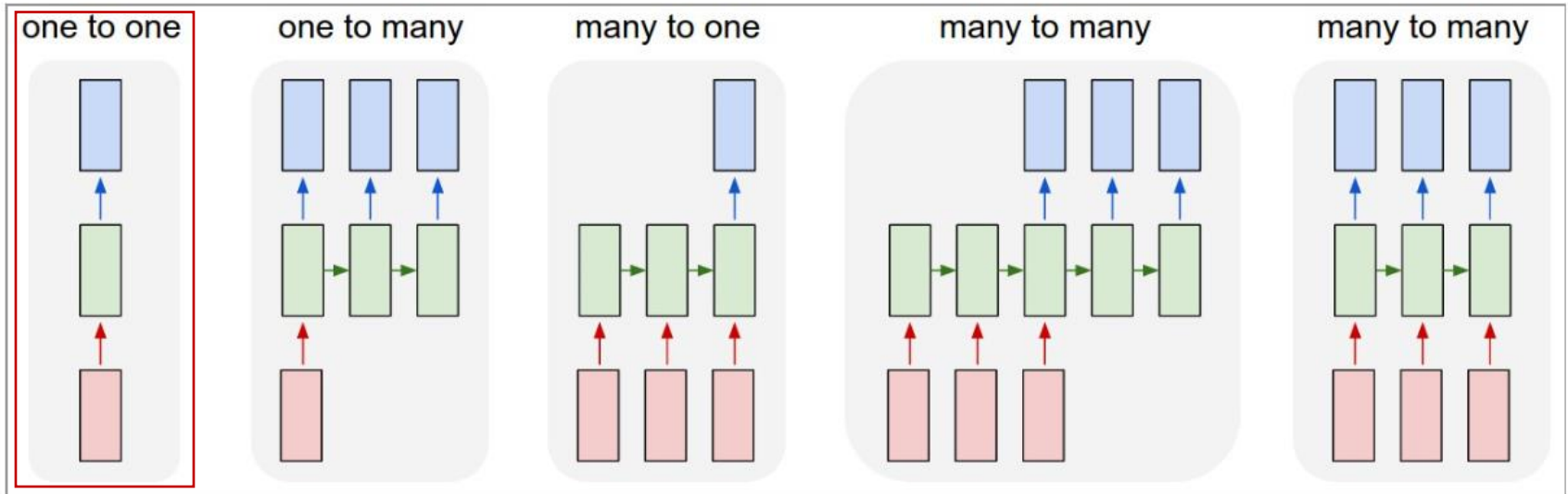# Sequential Data

- Sequences in our world:
    - Audio
    - Text
    - Video
    - Weather
    - Stock market

- RNNs are tools for making predictions about sequences.
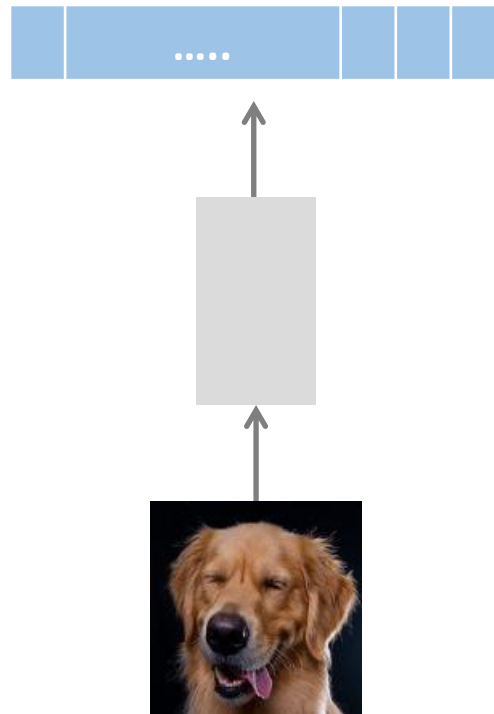
Recurrent Neural Networks

one to one | one to many | many to one | many to many | many to many

Vanilla NN

Recurrent NNs

One-to-one

Vanilla mode of processing without RNN

Example: Image classification

Example: One-to-one

Vanilla mode of processing without RNN
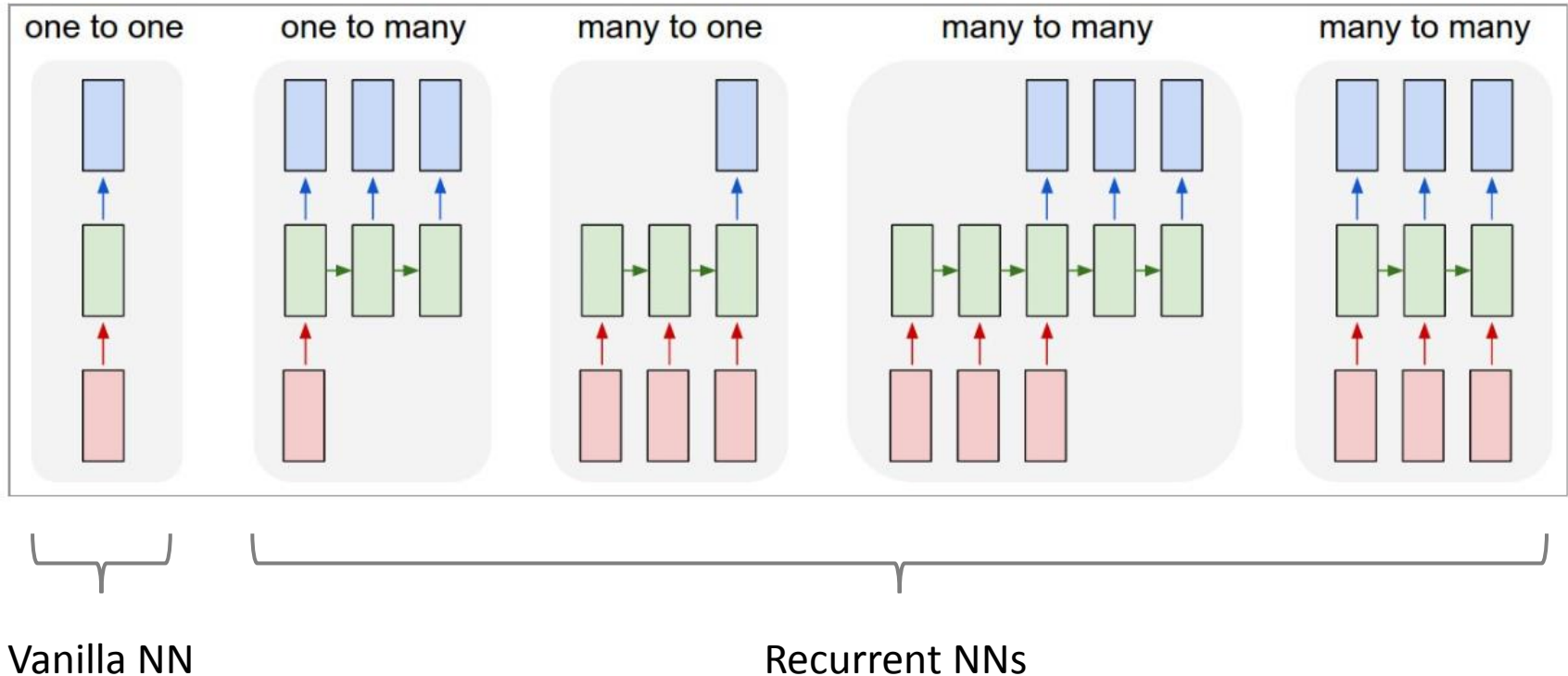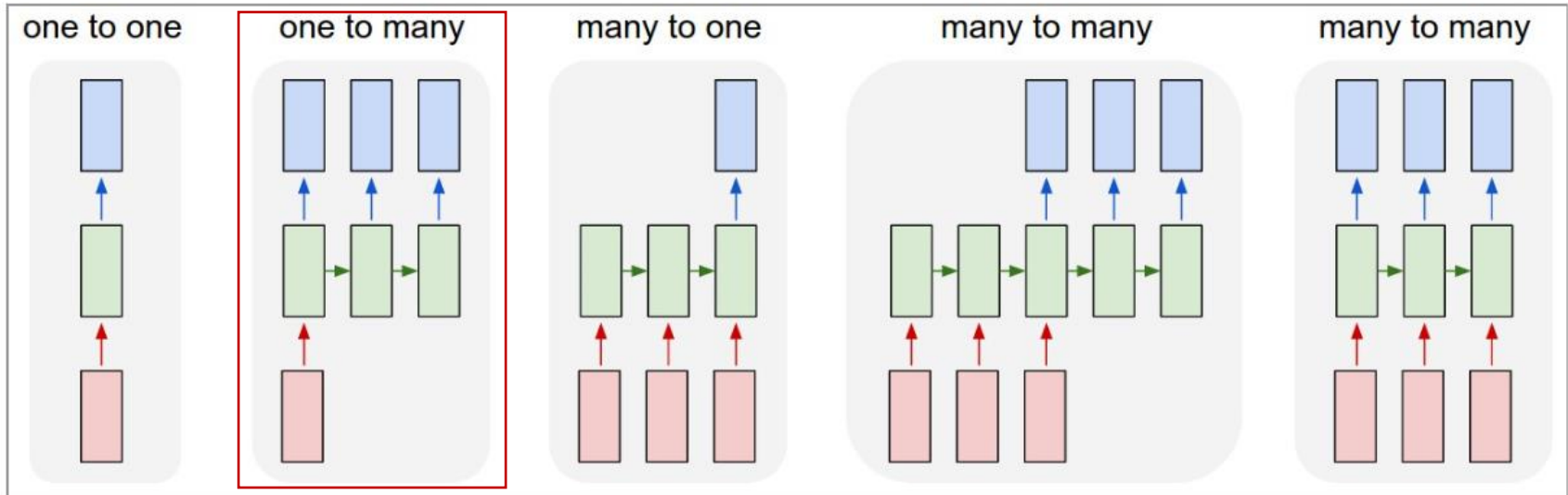
Example: Image classification

Prediction: "Dog"
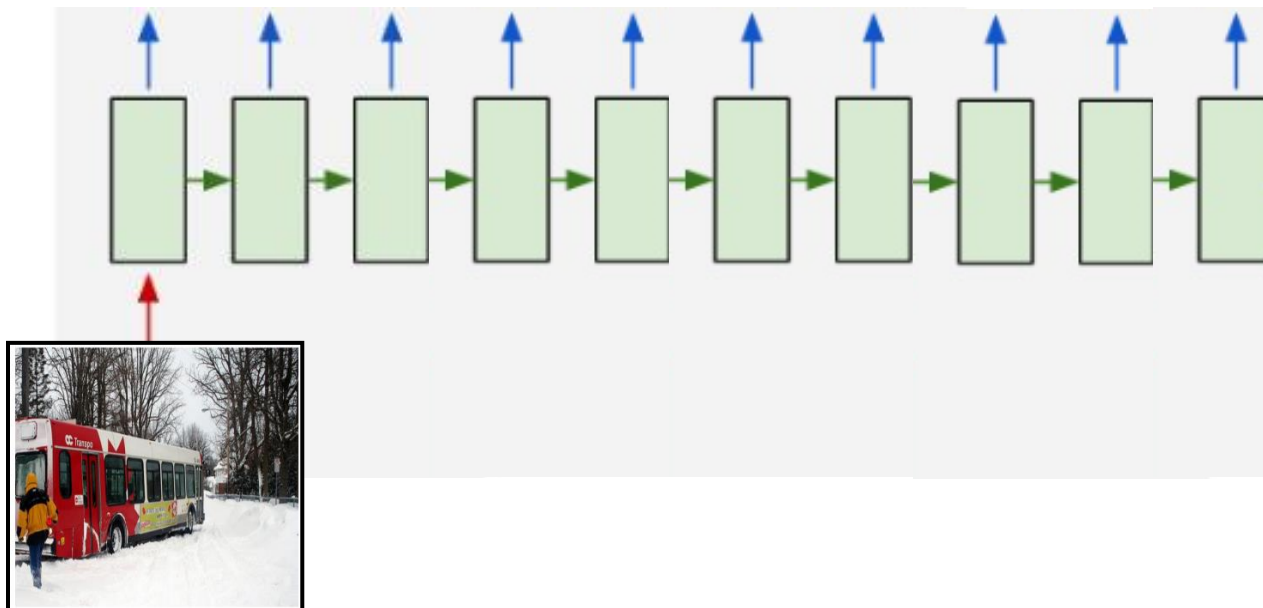
Vanilla NN

Image

Recurrent Neural Networks

Vanilla NN

Recurrent NNs

One-to-many

Sequence output

Example: Image captioning

Example: One-to-many

Sequence output

Example: Image Captioning

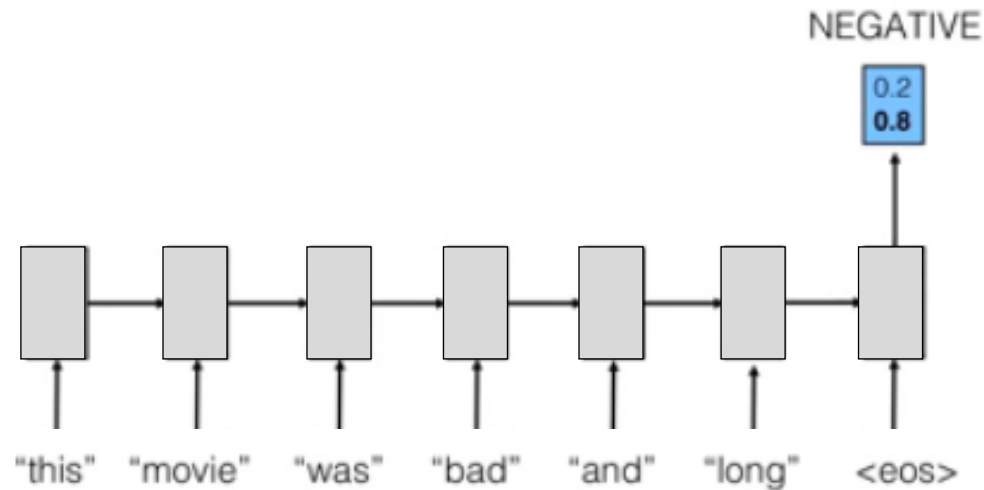Venugopalan et al., arXiv, https://arxiv.org/pdf/1606.07770.pdf
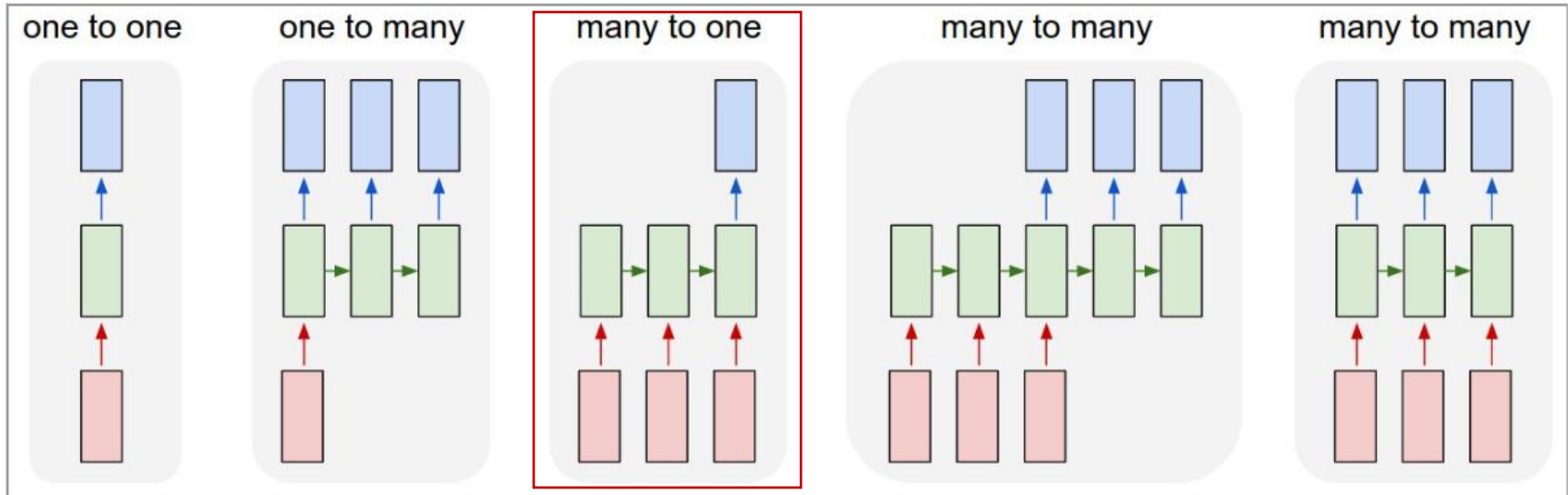
Many-to-one

Sequence input

Examples: Sentiment analysis
Action recognition

Example: Many-to-one

Sequence input

Example: Sentiment analysis

NEGATIVE

| 0.2 |
| **0.8** |

"this" "movie" "was" "bad" "and" "long" <eos>
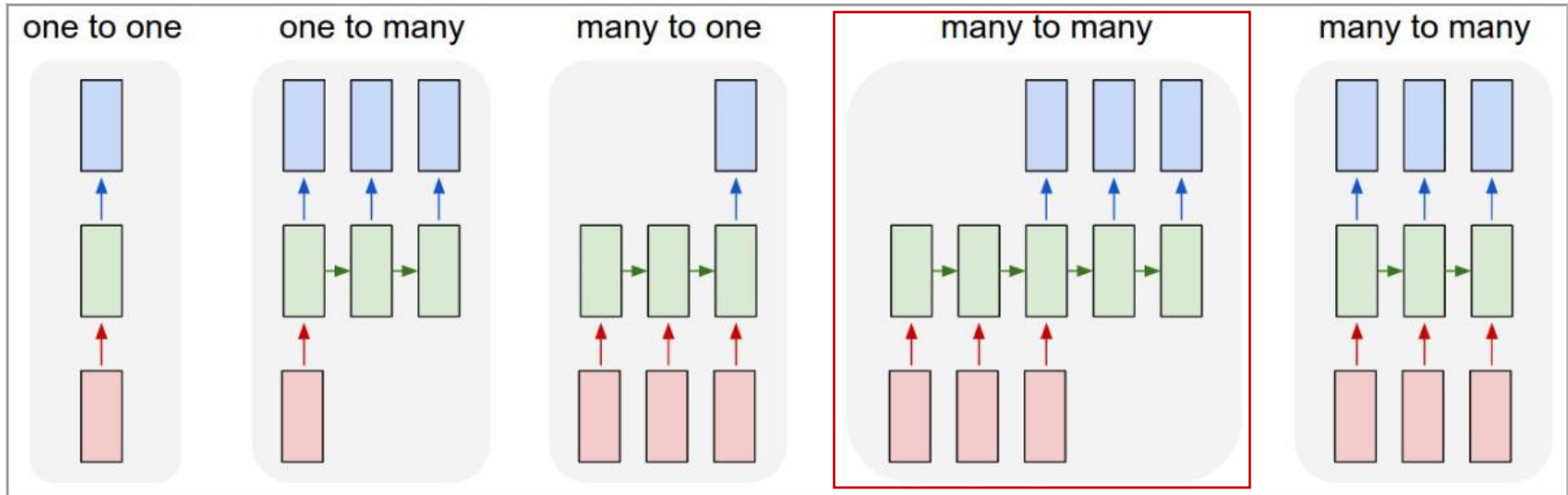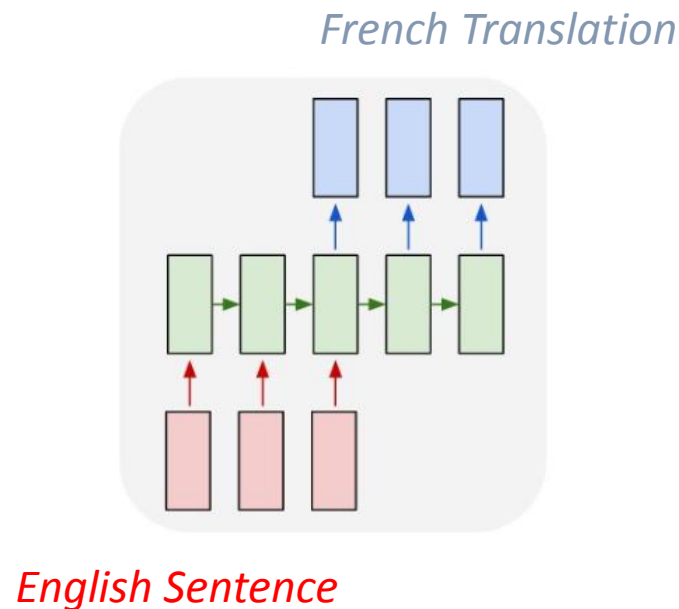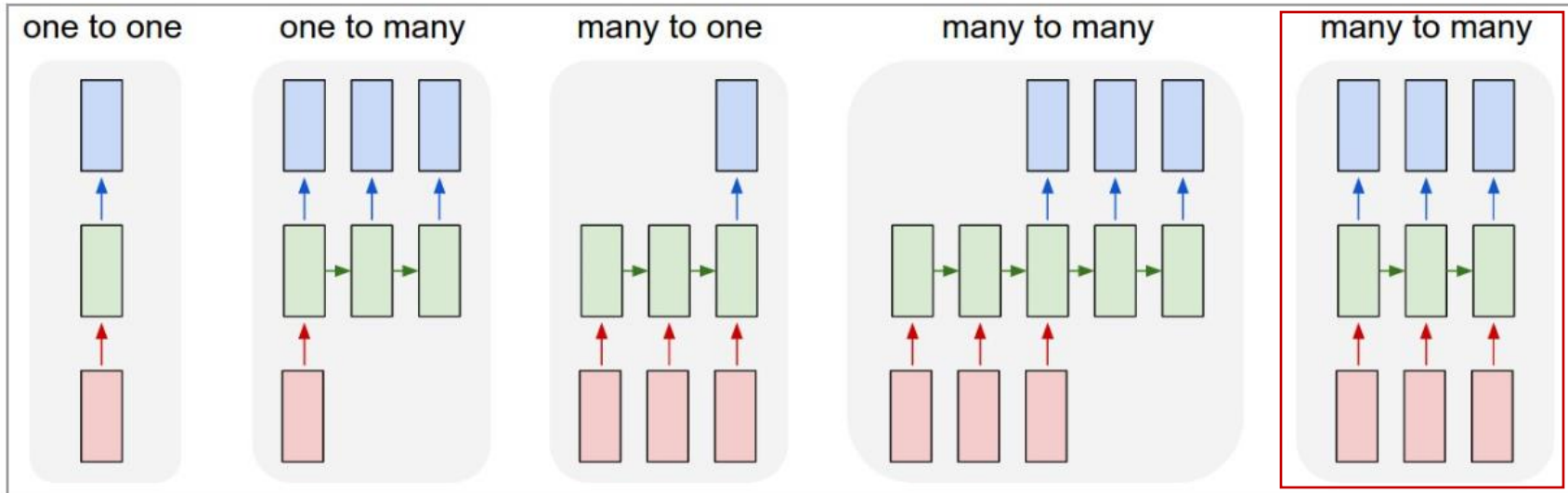
Many-to-many

Sequence input and sequence output

Example: Machine translation

Example: Many-to-many

Sequence input and sequence output

Example: Machine translation

French Translation

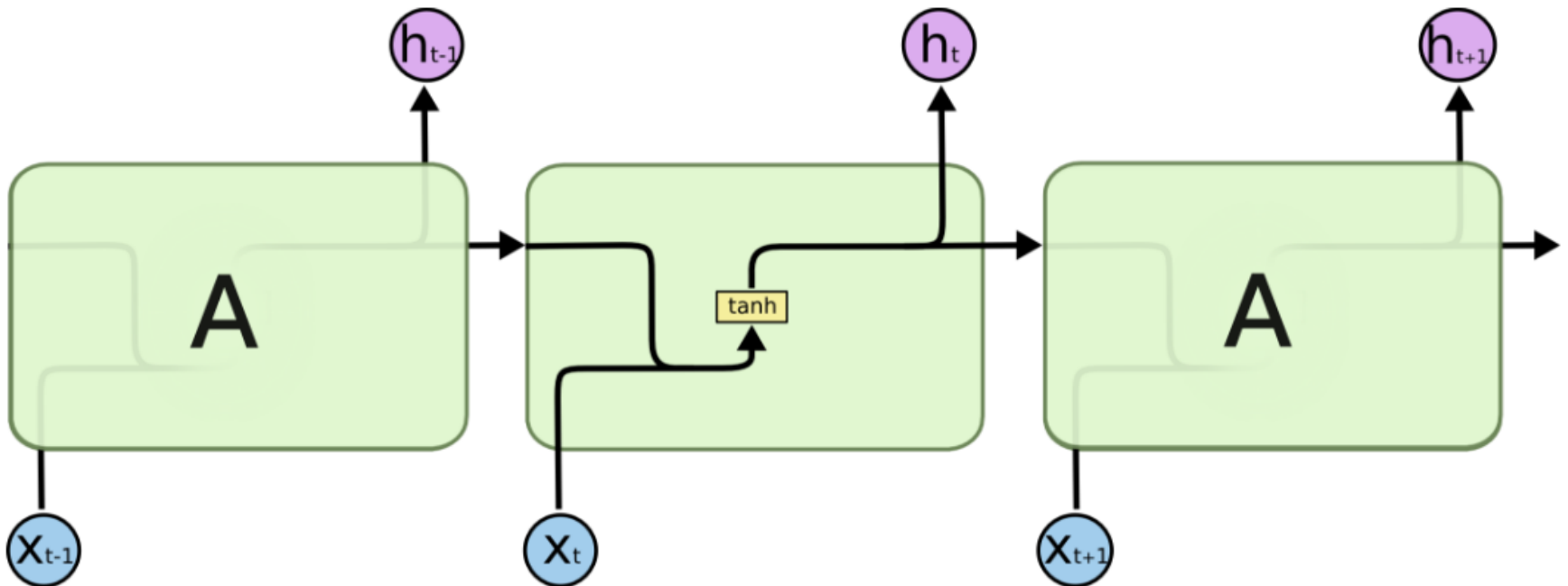*English Sentence*

Synced Many-to-many
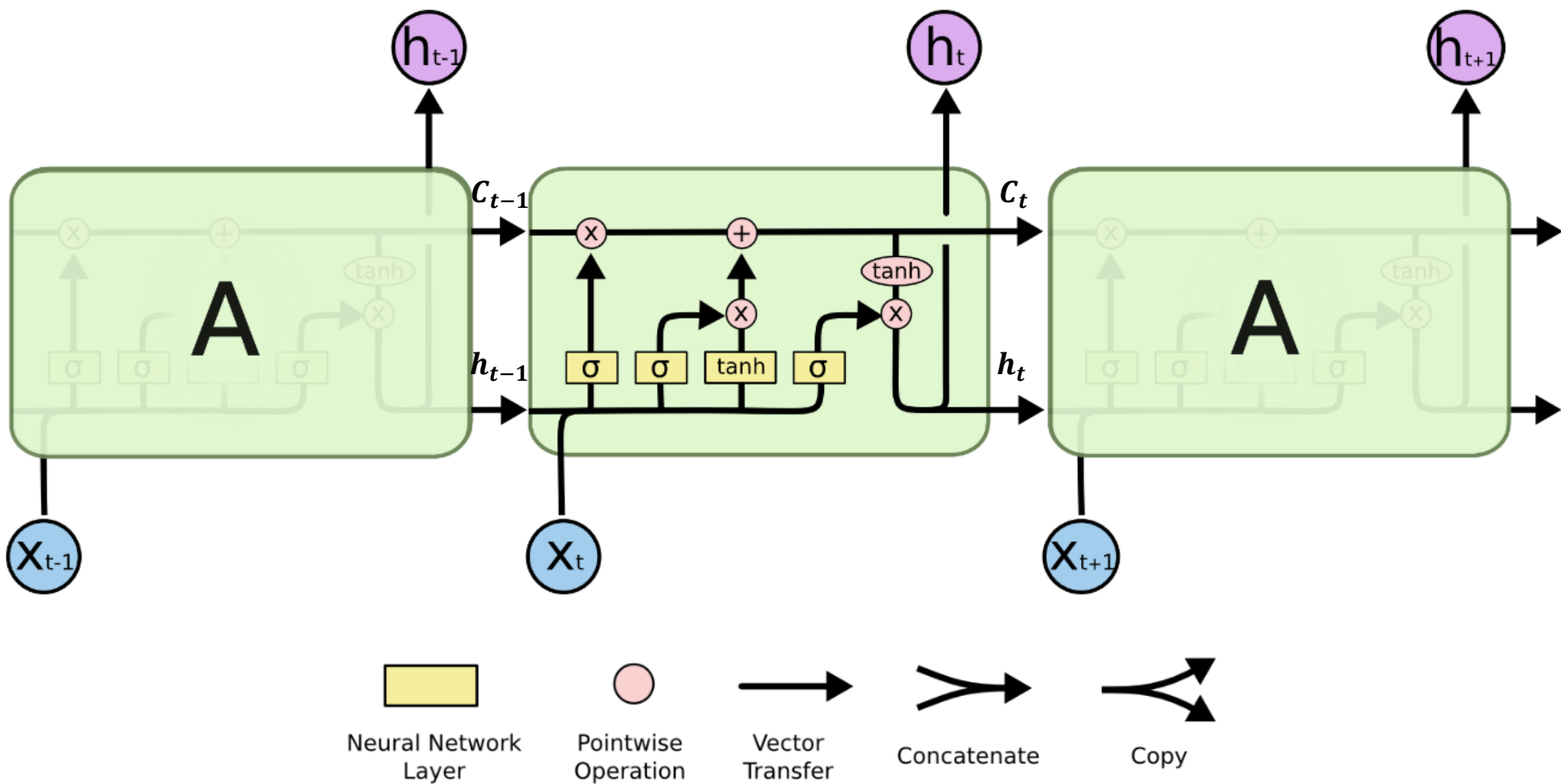

Synced sequence input and output

Examples: Tracking

      Early action detection

RNNs

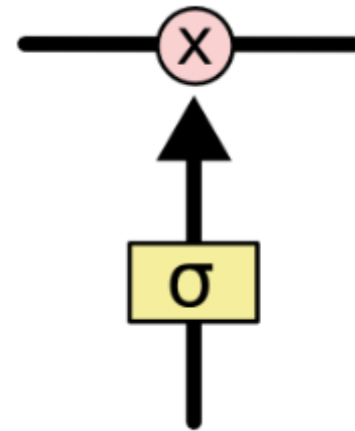- In a standard RNN the repeating module has a simple structure. Example:

Chris Olah, Blog, http://colah.github.io/posts/2015-08-Understanding-LSTMs/

21

LSTMs

Chris Olah, Blog, http://colah.github.io/posts/2015-08-Understanding-LSTMs/

LSTM Memory / Cell State

C_{t-1}

Cell-state @t-1

h_t

C_t

Cell-state @t

tanh

f_t i_t o_t

σ   σ   tanh   σ

Cell

h_{t-1}

h_t

x_t

Chris Olah, Blog, http://colah.github.io/posts/2015-08-Understanding-LSTMs/
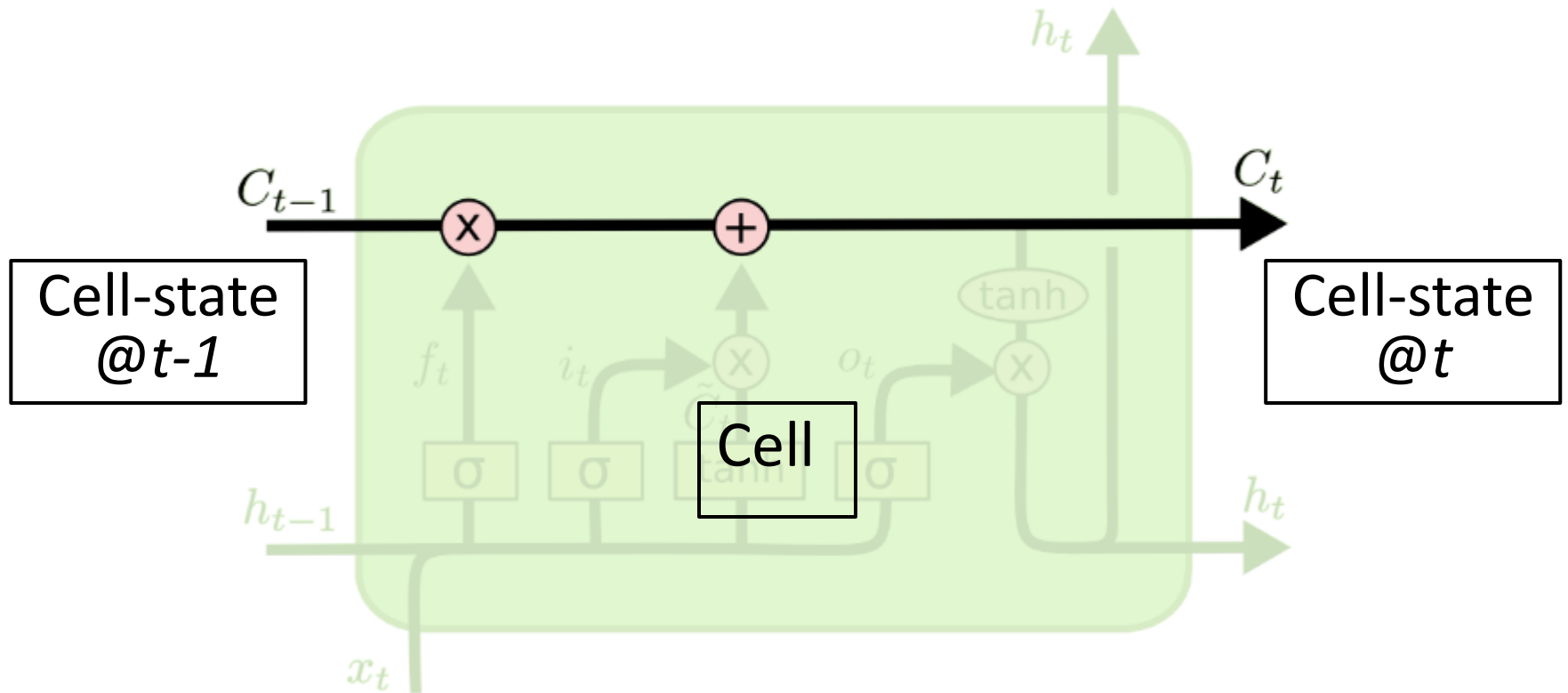
Gate

- Composed of a sigmoid neural net layer and a pointwise multiplication operation.

Chris Olah, Blog, http://colah.github.io/posts/2015-08-Understanding-LSTMs/

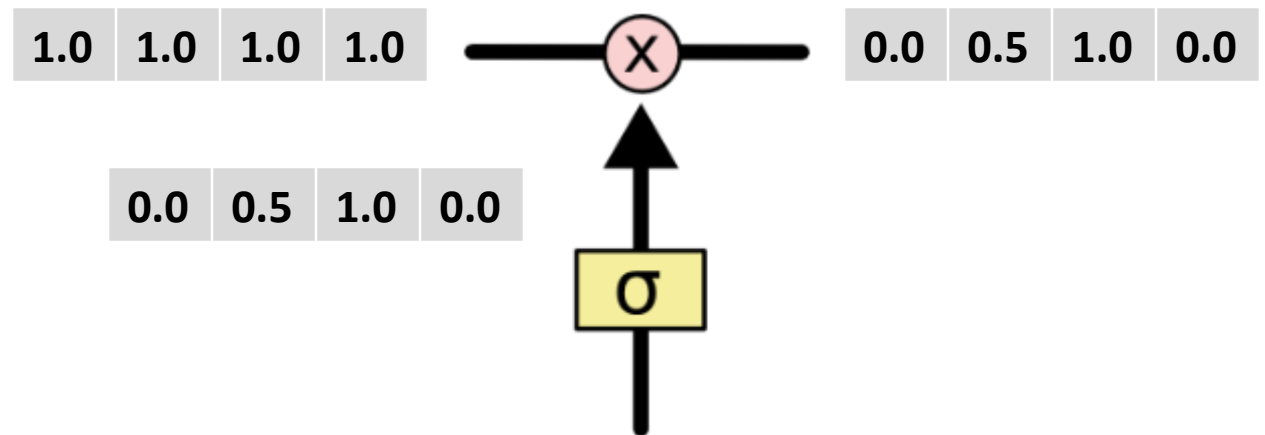# Gate

- sigmoid: outputs numbers between:
  - zero "let nothing through," and
  - one, "let everything through!"
- Example:

gate input ——⊗—— gate output

σ

Chris Olah, Blog, http://colah.github.io/posts/2015-08-Understanding-LSTMs/
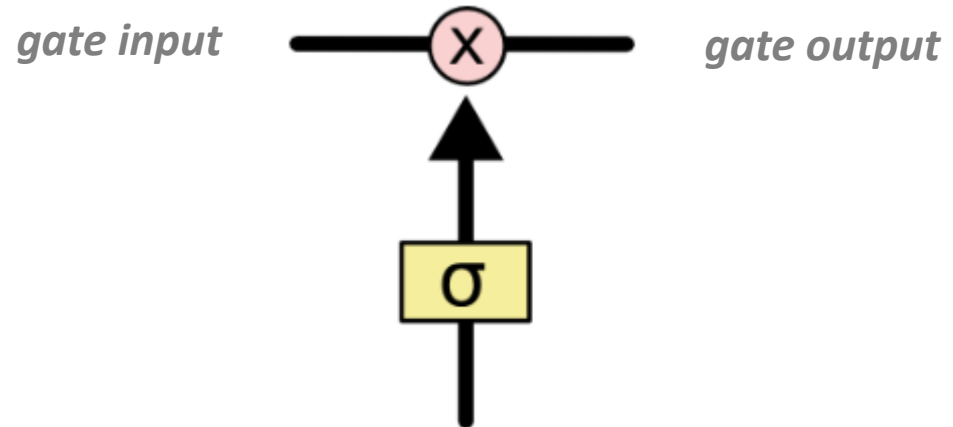
Gate

- sigmoid: outputs numbers between:
  - zero "let nothing through," and
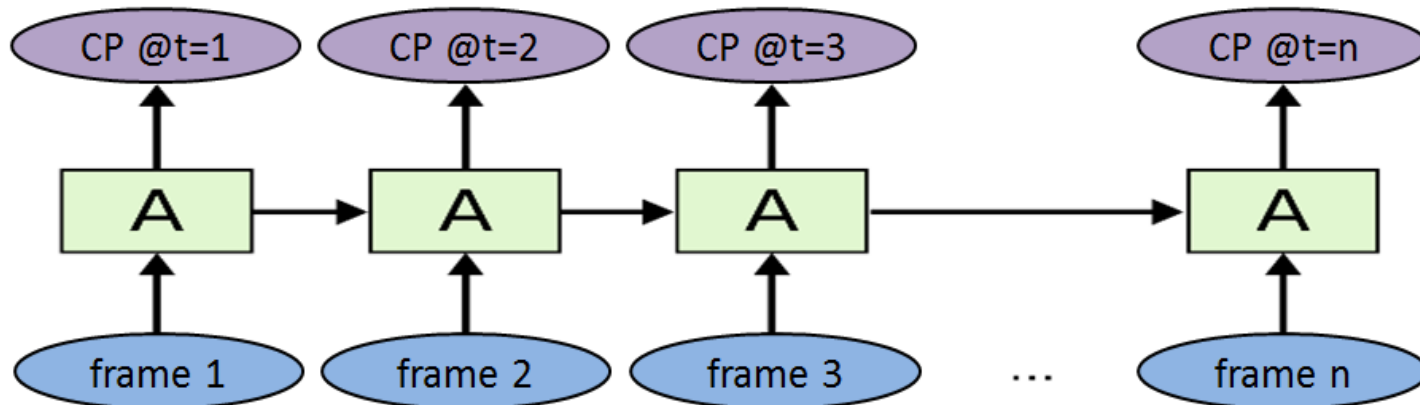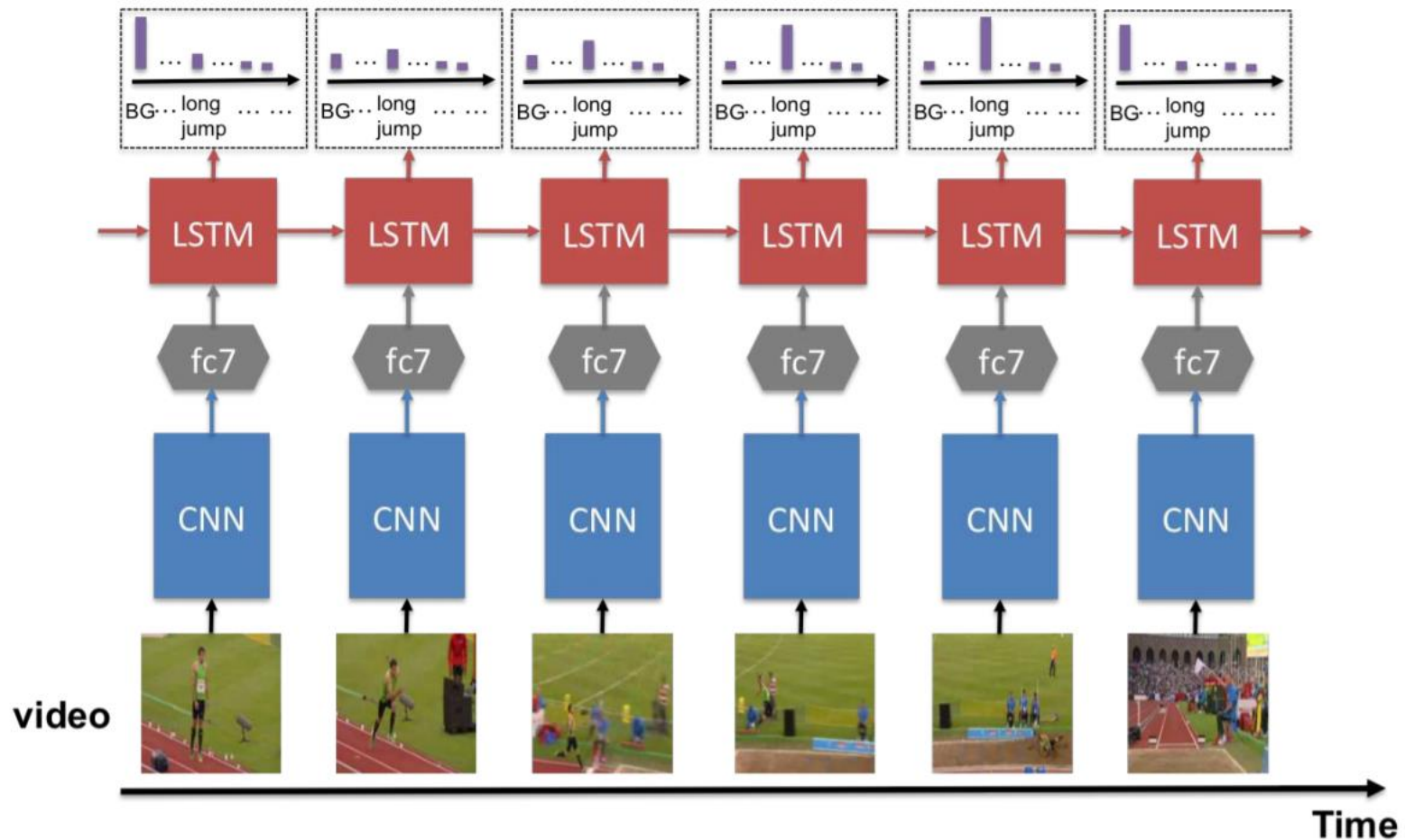 - one, "let everything through!"
- Example:

Neural Networks VI

Applications of Recurrent Networks

Application 1: Video Classification

- CP: conditional class probability
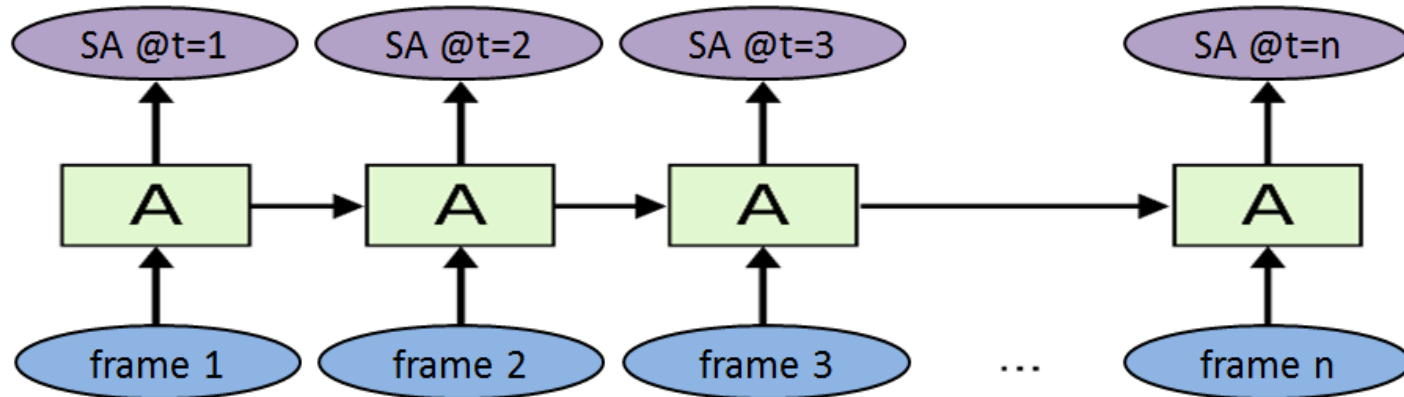- (frame i) could be a feature describing frame i, example: CNN feature

Application 1: Video Classification

Shugao Ma et al., CVPR'16, http://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/
Ma_Learning_Activity_Progression_CVPR_2016_paper.pdf

Application 2: Self-Driving Cars

- SA: steering angle
- frame i could be a feature describing frame i, example: 3D-CNN feature

Application 2: Self-Driving Cars

- DeepTesla

DeepTesla, http://selfdrivingcars.mit.edu/deeptesla/

Application 2: Self-Driving Cars

- Udacity winning team: *Team Komanda*
  - x_t: 3D convolution of image sequence
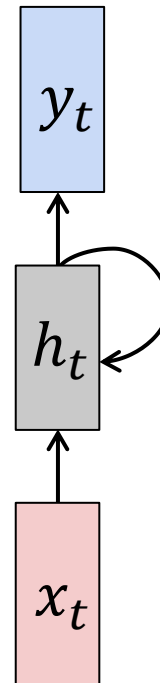 - h_t: steering angle, speed, torque

Actual = -0.55851, Predicted = -0.67336

Application 3: Character RNN

Character-level
language model
example

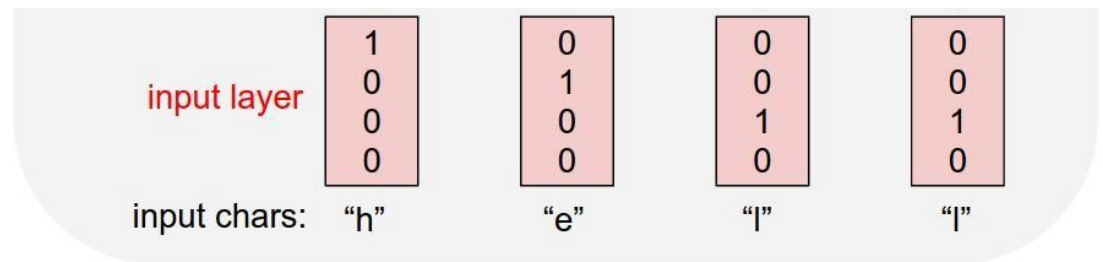Vocabulary:
[h,e,l,o]

Example training
sequence:
"hello"

$$y_t$$

$$h_t$$

$$x_t$$

Application 3: Character RNN

Character-level language model example

Vocabulary:
[h,e,l,o]

Example training sequence:
"hello"

Application 3: Character RNN

Character-level language model example

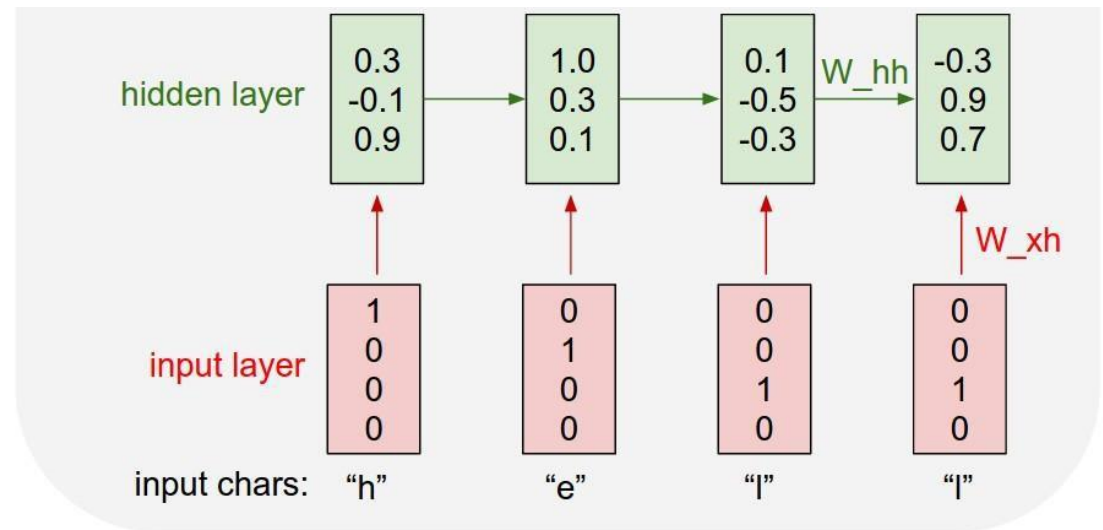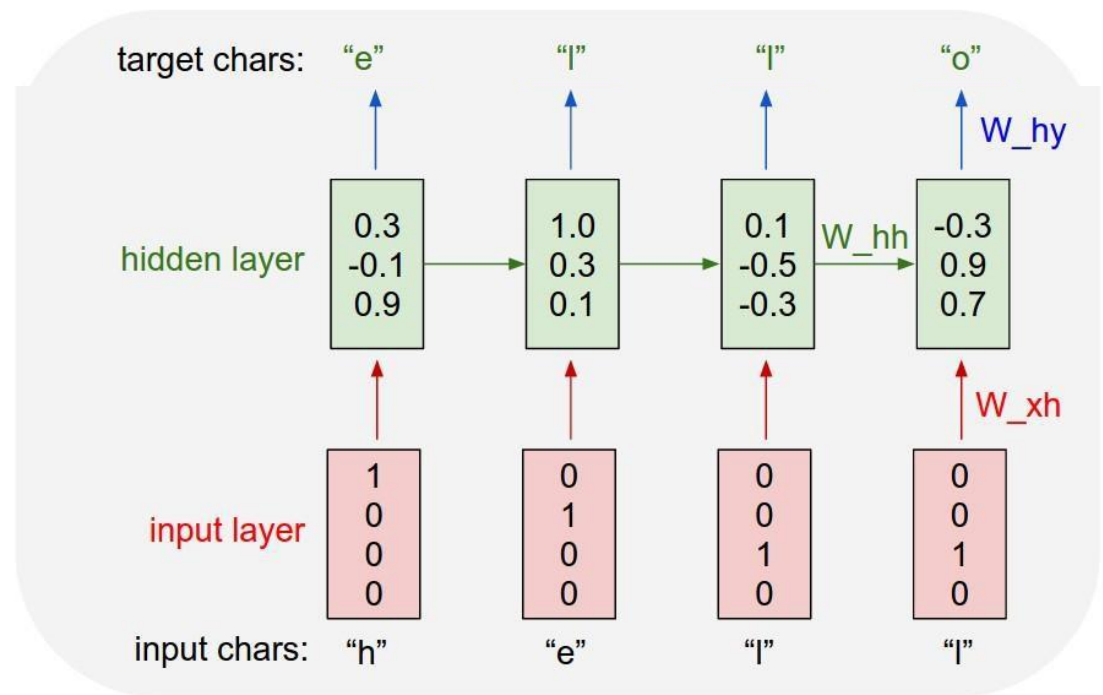Vocabulary: [h,e,l,o]

Example training sequence: **"hello"**

Application 3: Character RNN

Character-level language model example

Vocabulary: [h,e,l,o]

Example training sequence: **"hello"**

Application 4:Reading cursive

handwriting

- This is a natural task for an RNN.

- The input is a sequence of (x,y,p) coordinates of the tip of the pen, where p indicates whether the pen is up or down.

- The output is a sequence of characters.

- Graves & Schmidhuber (2009) showed that RNNs with LSTM are currently the best systems for reading cursive writing.

  - They used a sequence of small images as input rather than pen coordinates.

Application 5: StyleText Generation

Training text: William Shakespeare
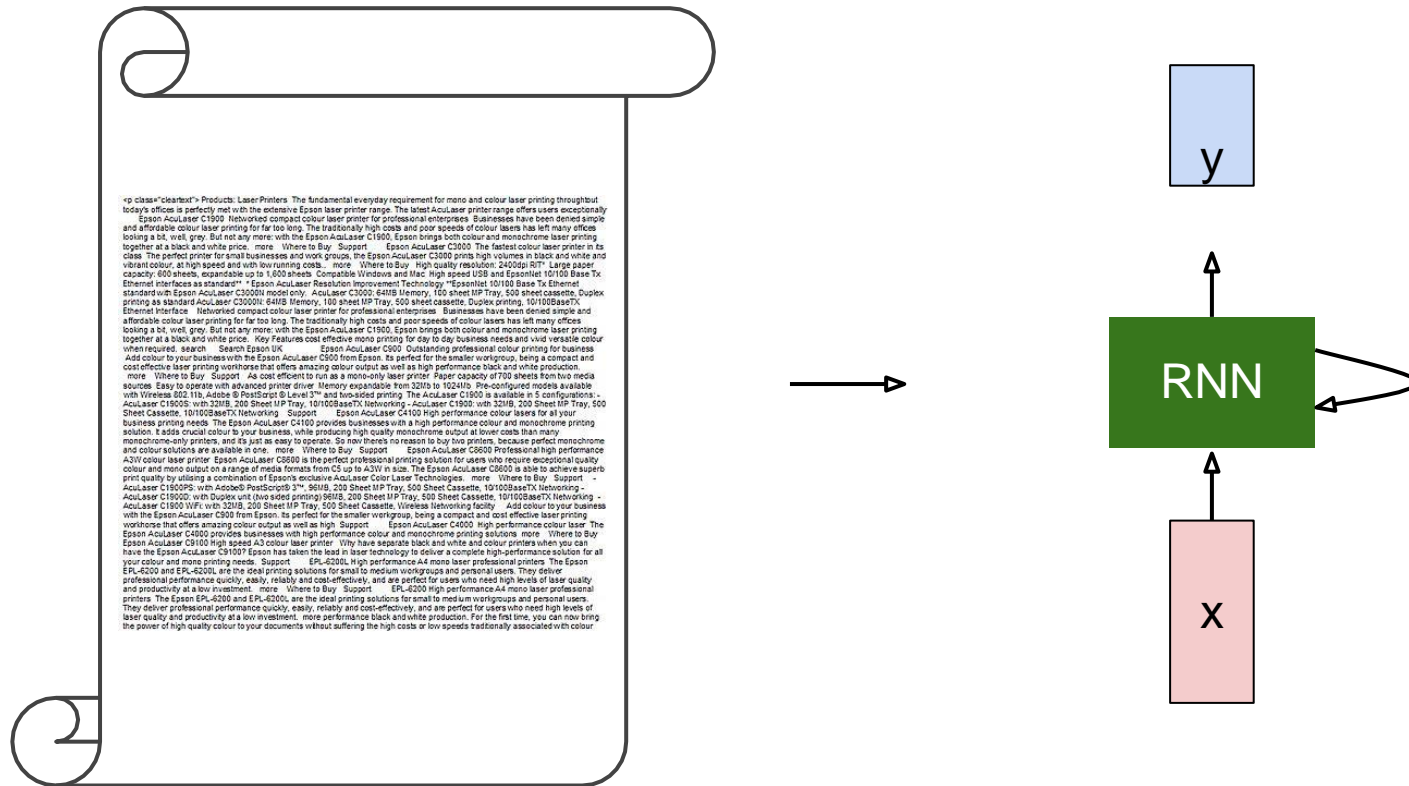
Application 5: StyleText Generation

at first:

```
tyntd-iafhatawiaoihrdemot  lytdws  e ,tfti, astai f ogoh eoase rrranbyne 'nhthnee e
plia tklrgd t o idoe ns,smtt   h ne etie h,hregtrs nigtike,aoaenns lng
```

↓ train more

```
"Tmont thithey" fomesscerliund
Keushey. Thom here
sheulke, anmerenith ol sivh I lalterthend Bleipile shuwy fil on aseterlome
coaniogennc Phe lism thond hon at. MeiDimorotion in ther thize."
```

↓ train more

```
Aftair fall unsuch that the hall for Prince Velzonski's that me of
her hearly, and behs to so arwage fiving were to it beloge, pavu say falling misfort
how, and Gogition is so overelical and ofter.
```

↓ train more

```
"Why do what that day," replied Natasha, and wishing to himself the fact the
princess, Princess Mary was easier, fed in had oftened him.
Pierre aking his soul came to the packs and drove up his father-in-law women.
```

Application 6: Code Generation

Train on C code

Application 6: Code Generation

```c
static void do_command(struct seq_file *m, void *v)
{
  int column = 32 << (cmd[2] & 0x80);
  if (state)
    cmd = (int)(int_state ^ (in_8(&ch->ch_flags) & Cmd) ? 2 : 1);
  else
    seq = 1;
  for (i = 0; i < 16; i++) {
    if (k & (1 << 1))
      pipe = (in_use & UMXTHREAD_UNCCA) +
        ((count & 0x00000000ffffff8) & 0x000000f) << 8;
    if (count == 0)
      sub(pid, ppc_md.kexec_handle, 0x20000000);
    pipe_set_bytes(i, 0);
  }
  /* Free our user pages pointer to place camera if all dash */
  subsystem_info = &of_changes[PAGE_SIZE];
  rek_controls(offset, idx, &soffset);
  /* Now we want to deliberately put it to device */
  control_check_polarity(&context, val, 0);
  for (i = 0; i < COUNTER; i++)
    seq_puts(s, "policy ");
}
```

Generated
C code

Application 7: Writing a Movie Script

https://arstechnica.com/the-multiverse/2016/06/an-ai-wrote-this-movie-and-its-strangely-moving/