Practical Machine Learning

Recurrent Neural Network

Today we cover

Review RNNs

RNN for language processing

Auto Encoders

Generate images

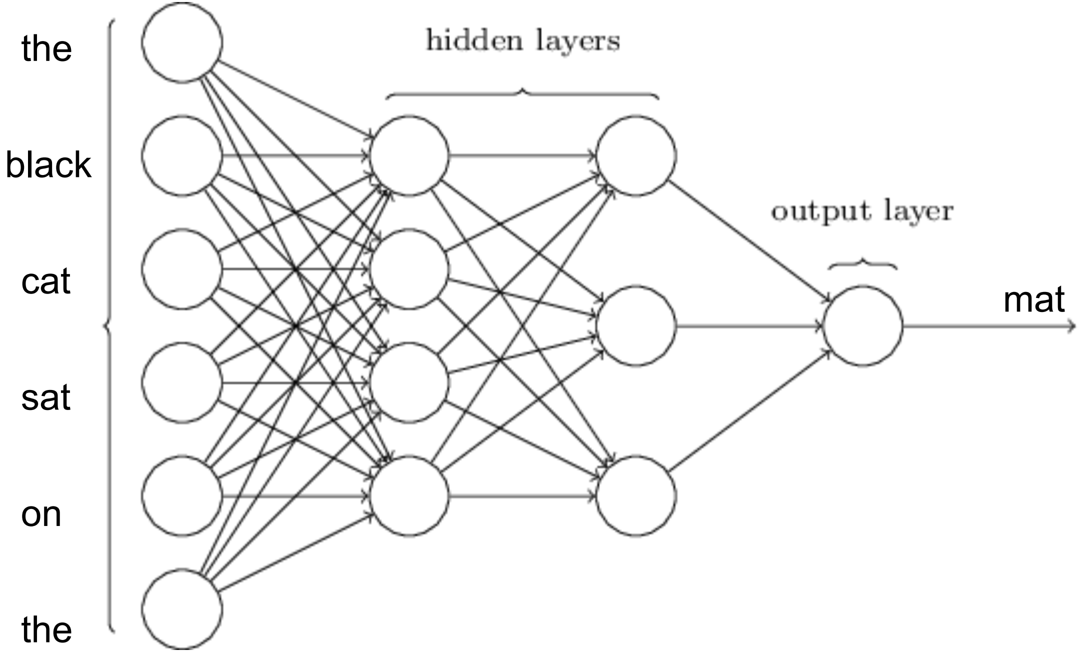
Recurrent Neural Networks

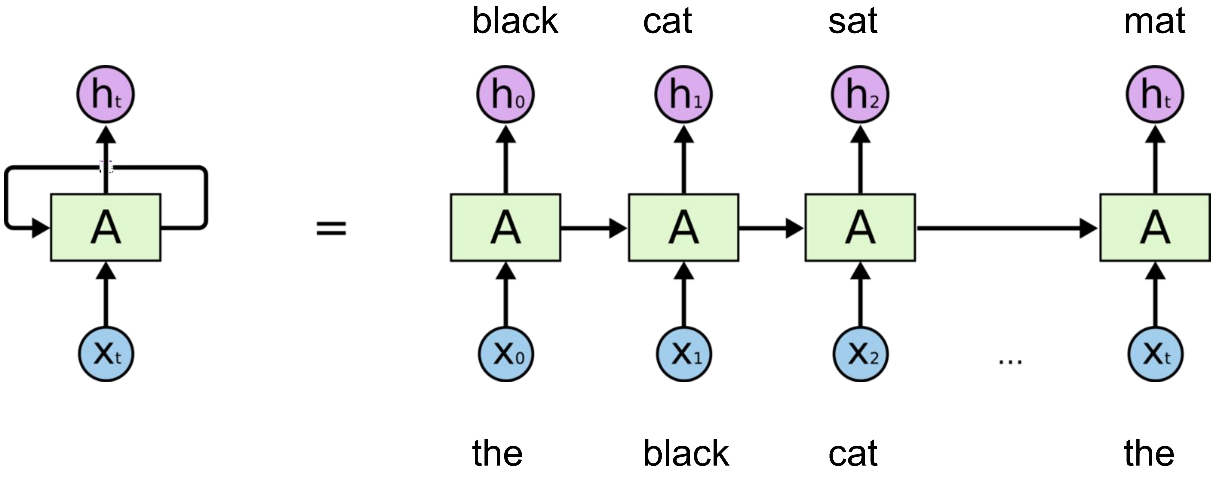
When to use them?

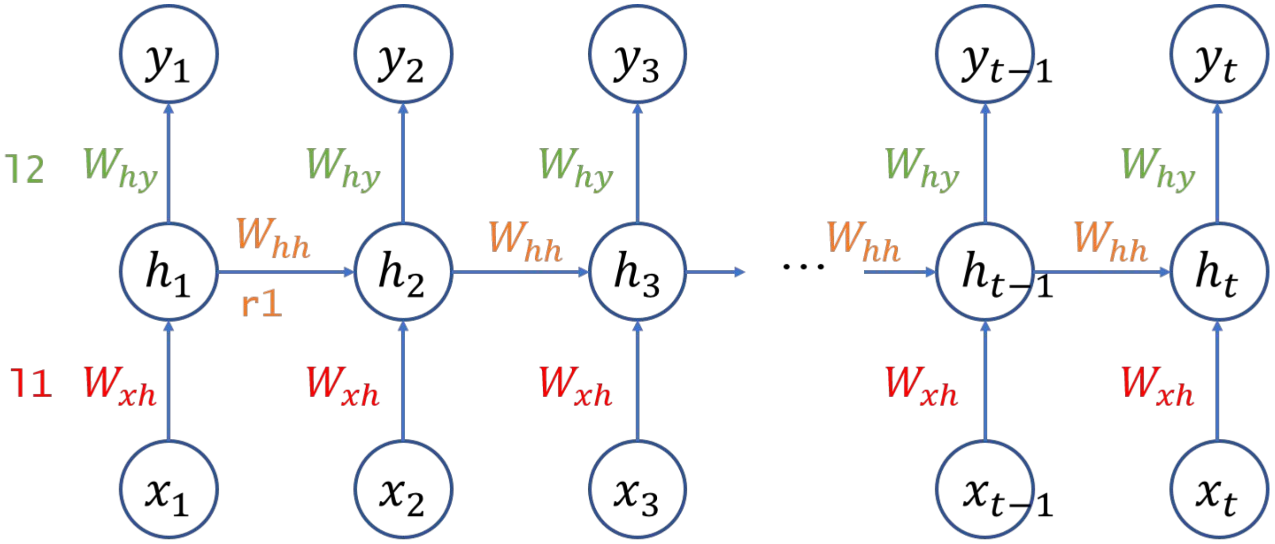
- I have a sequence
- Length unknown in advance

Examples:

- Time series
- Text
- Music


the
black
cat
sat
on
the
hidden layers
output layer
mat

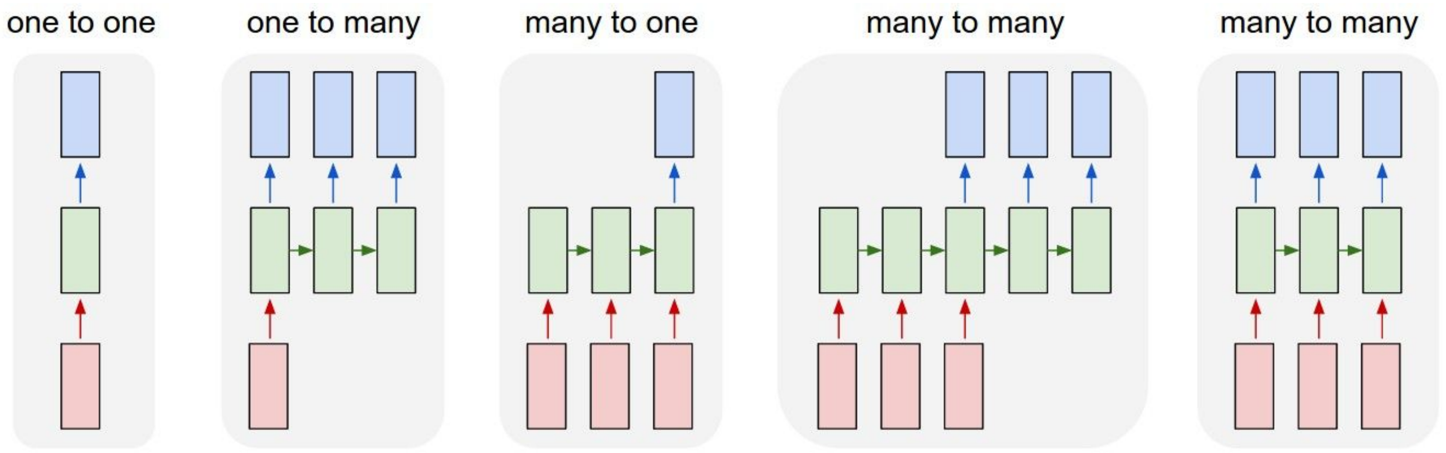

black
cat
sat
mat
h_t
h_0
h_1
h_2
h_t
A
=
A
A
A
A
x_t
x_0
x_1
x_2
...
x_t
the
black
cat
the



$$h_t = \tanh(\ l1(x_t) + r1(h_{t-1})\ )$$

$$y_t = l2(h_t)$$

Flavors of RNNs



Classifier

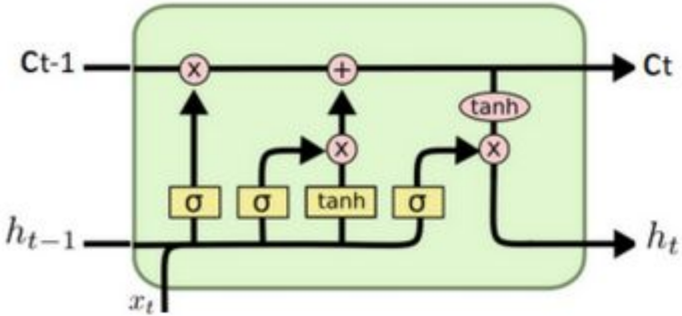
Generate text

Sentiment analysis

Translation

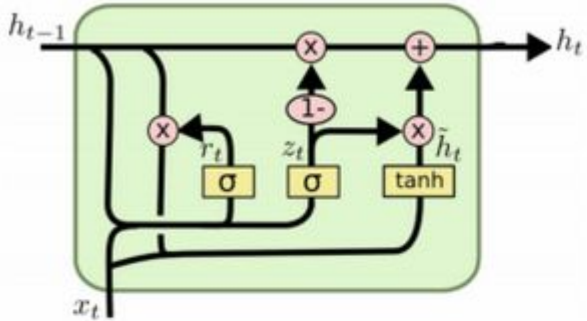
Time series prediction

LSTM



Ct-1
Ct
h_{t-1}
h_t
x_t
x
+
tanh
x
x
σ
σ
tanh
σ

GRU



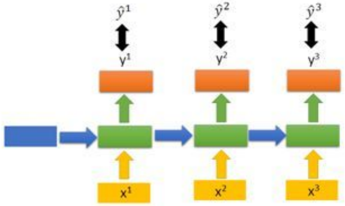
h_{t-1}
h_t
x
+
1-
x
x
r_t
z_t
$\tilde{h}_t$
σ
σ
tanh
x_t

Vanilla RNN

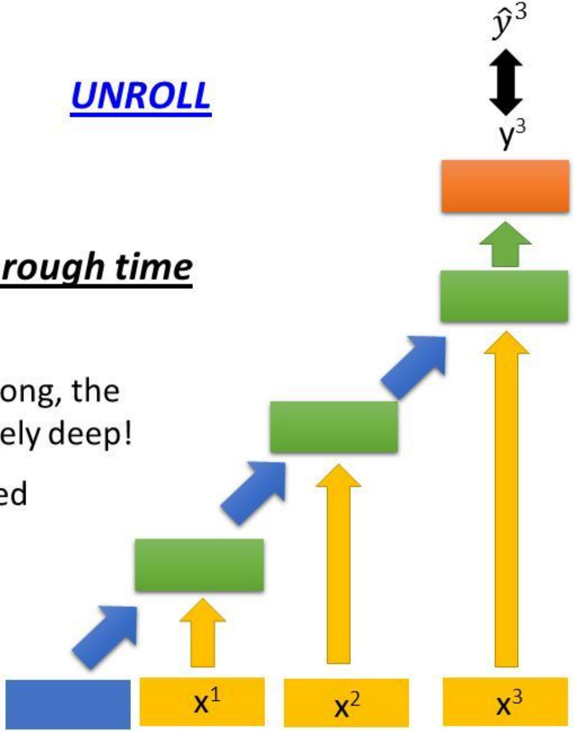
$$\boxed{h_t = \tanh\left(W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix}\right)}$$

LSTM

$$\boxed{\begin{gathered} \begin{pmatrix} i \\ f \\ o \\ g \end{pmatrix} = \begin{pmatrix} \sigma \\ \sigma \\ \sigma \\ \tanh \end{pmatrix} W \begin{pmatrix} h_{t-1} \\ x_t \end{pmatrix} \\ c_t = f \odot c_{t-1} + i \odot g \\ h_t = o \odot \tanh(c_t) \end{gathered}}$$



UNROLL



Backpropagation through time (BPTT)

When the sequence is long, the network can be extremely deep!

Some weights are shared

One hot encoding

Encode: A, B, C, D

A => [1, 0, 0, 0]

B => [0, 1, 0, 0]

C => [0, 0, 1, 0]

D => [0, 0, 0, 1]

Data preparation

```
X = [

        [[0, 0, 0, 1], [1, 0, 0, 0] …..],

        … ]

Y = [  [0, 1, 0, 0],

        … ]
```

Code

```
model = Sequential()
model.add(LSTM(128, input_shape=(lstm_size, len(char_dict))))
model.add(Dense(units=len(char_dict), activation='softmax'))
…..
```

You code

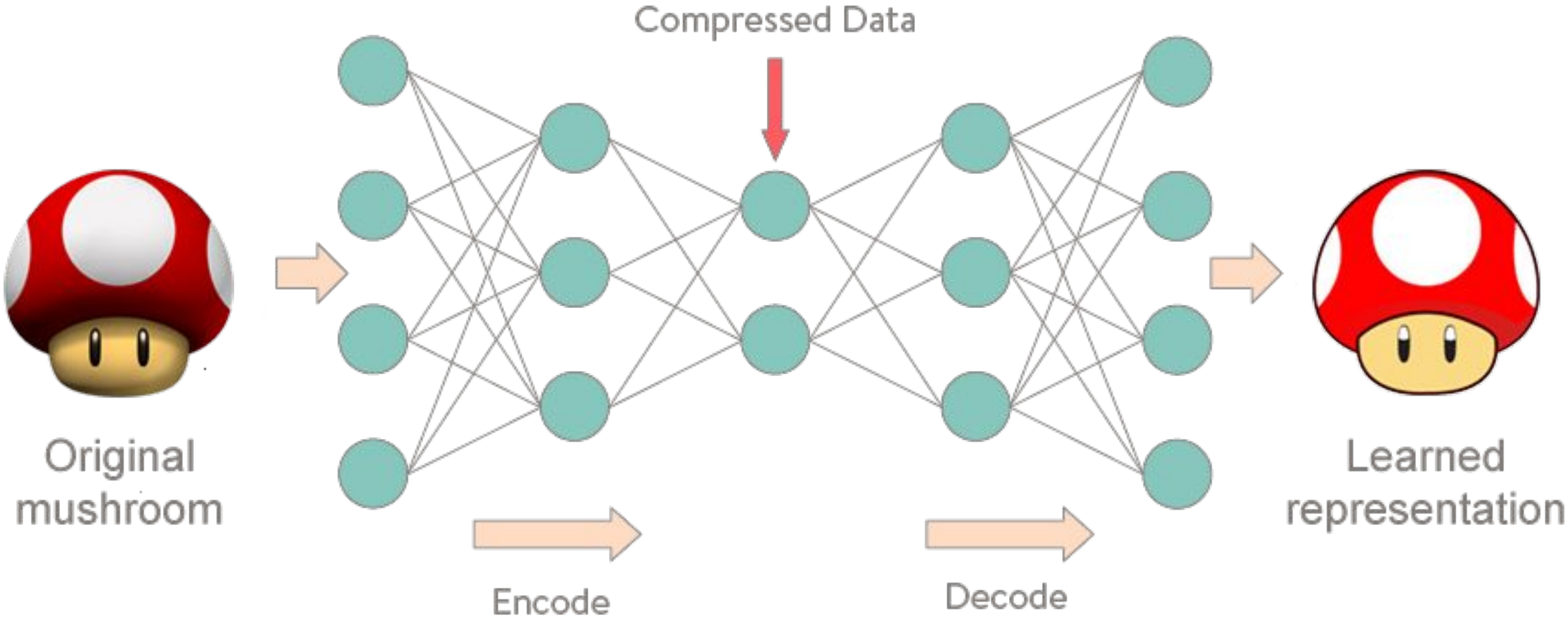
Get text

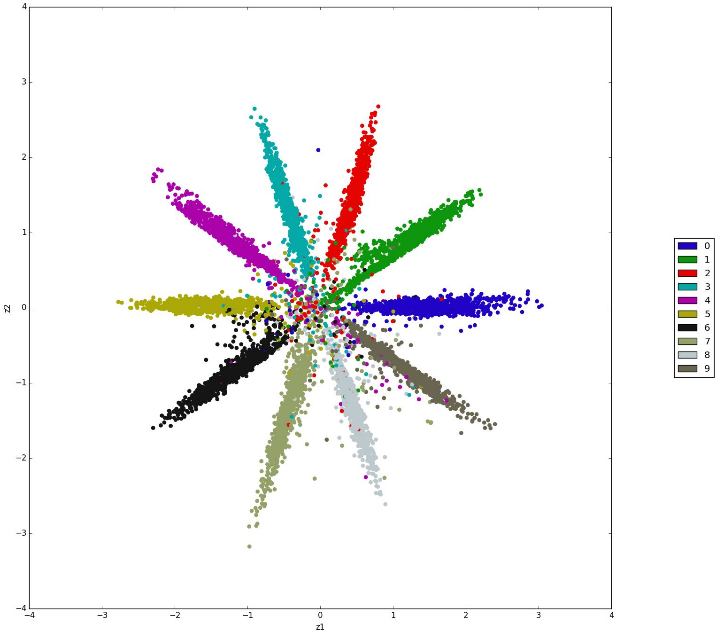
Build next word predictor

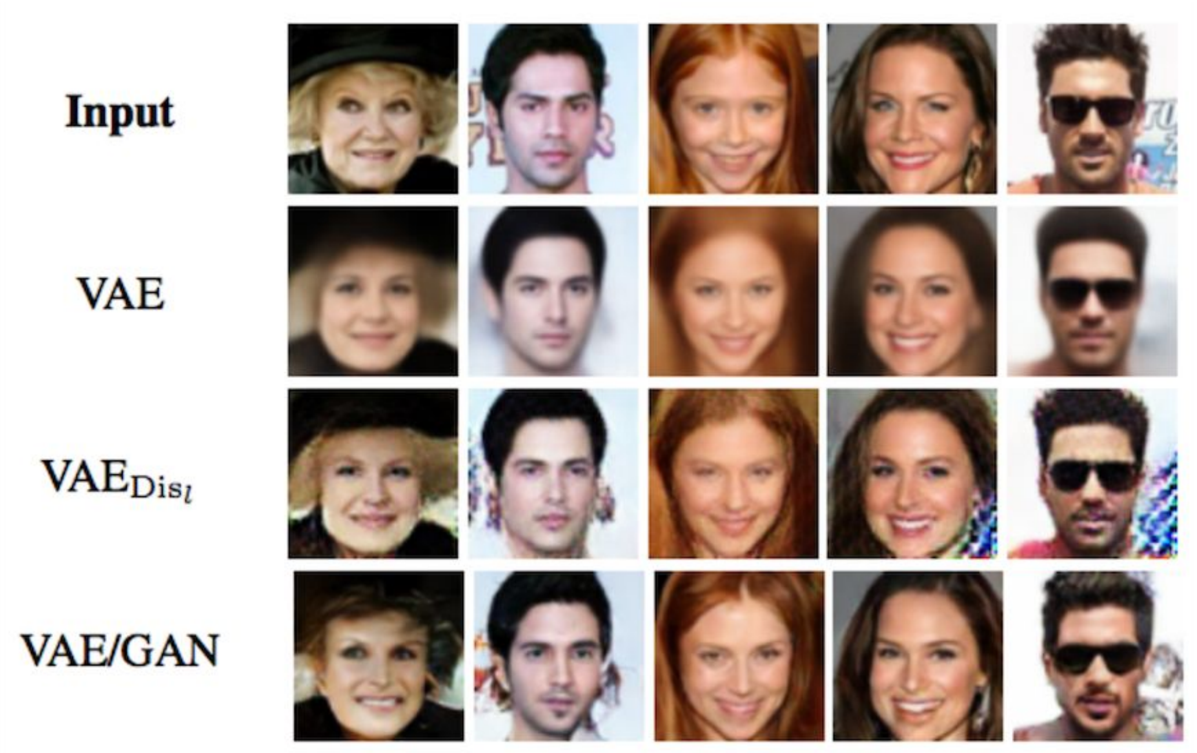
Autoencoders

Learning representations

=> [0.3, 0.4, 0.7]


Compressed Data
Original mushroom
Encode
Decode
Learned representation


4
3
2
1
0
−1
−2
−3
−4
−4
−3
−2
−1
0
1
2
3
4
z1
z2
0
1
2
3
4
5
6
7
8
9


Input
VAE
$\mathrm{VAE}_{\mathrm{Dis}_l}$
VAE/GAN

Code

Fashion-MNIST autoencoder

https://blog.keras.io/building-autoencoders-in-keras.html