Artificial Intelligence in Finance and Quantitative Analysis

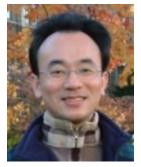


Deep Learning in Finance Reinforcement Learning in Finance

1111AIFQA09 MBA, IM, NTPU (M6132) (Fall 2022) Tue 2, 3, 4 (9:10-12:00) (B8F40)







Min-Yuh Day, Ph.D, Associate Professor

Institute of Information Management, National Taipei University

https://web.ntpu.edu.tw/~myday



Syllabus



Week Date Subject/Topics

- 1 2022/09/13 Introduction to Artificial Intelligence in Finance and Quantitative Analysis
- 2 2022/09/20 Al in FinTech: Metaverse, Web3, DeFi, NFT, Financial Services Innovation and Applications
- 3 2022/09/27 Investing Psychology and Behavioral Finance
- 4 2022/10/04 Event Studies in Finance
- 5 2022/10/11 Case Study on AI in Finance and Quantitative Analysis I
- 6 2022/10/18 Finance Theory

Syllabus



Week Date Subject/Topics

- 7 2022/10/25 Data-Driven Finance
- 8 2022/11/01 Midterm Project Report
- 9 2022/11/08 Financial Econometrics and Machine Learning
- 10 2022/11/15 Al-First Finance
- 11 2022/11/22 Deep Learning in Finance;
 Reinforcement Learning in Finance
- 12 2022/11/29 Case Study on AI in Finance and Quantitative Analysis II

Syllabus



Week Date Subject/Topics

- 13 2022/12/06 Industry Practices of AI in Finance and Quantitative Analysis
- 14 2022/12/13 Algorithmic Trading; Risk Management;
 Trading Bot and Event-Based Backtesting
- 15 2022/12/20 Final Project Report I
- 16 2022/12/27 Final Project Report II
- 17 2023/01/03 Self-learning
- 18 2023/01/10 Self-learning

Deep Learning in Finance Reinforcement Learning in Finance

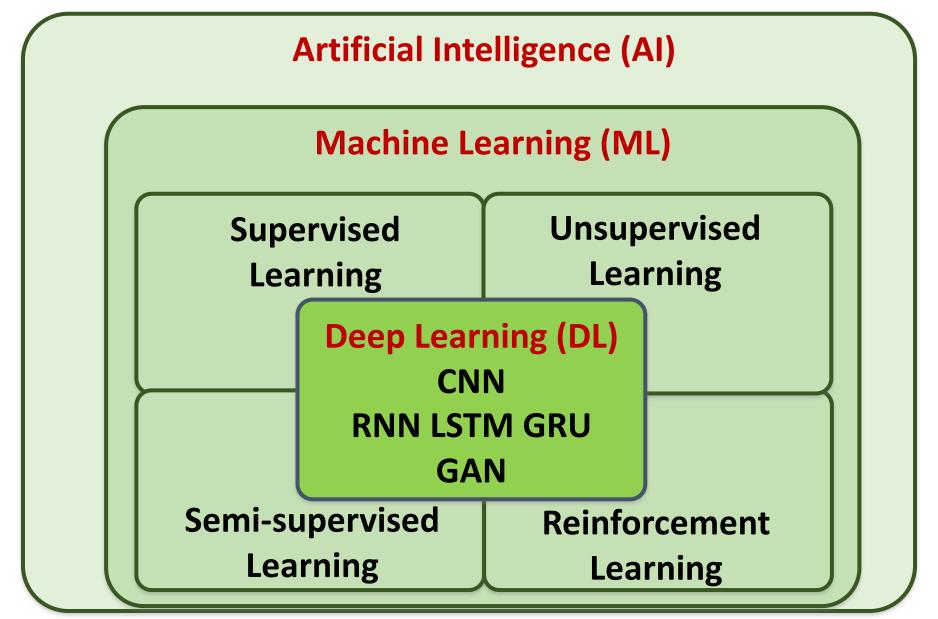
Outline

- Deep Learning (DL) in Finance
 - Dense Neural Networks (DNN)
 - Recurrent Neural Networks (RNN)
 - Convolutional Neural Networks (CNN)
- Reinforcement Learning (RL) in Finance
 - Q Learning (QL)
 - Improved Finance Environment
 - Improved Financial QL Agent

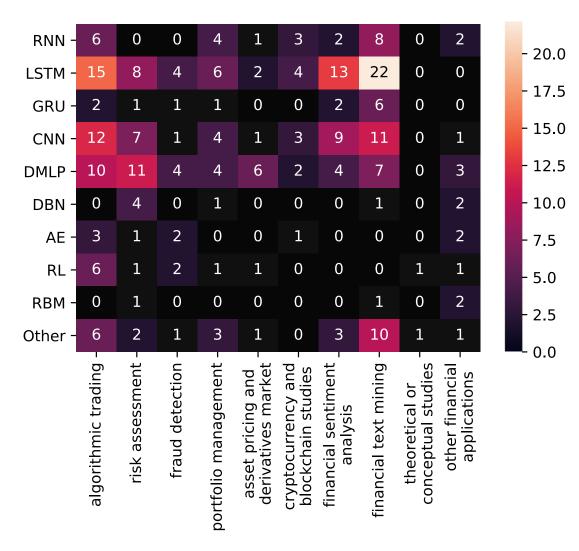
Deep Learning in Finance

- Dense Neural Networks (DNN)
- Recurrent Neural Networks (RNN)
- Convolutional Neural Networks (CNN)

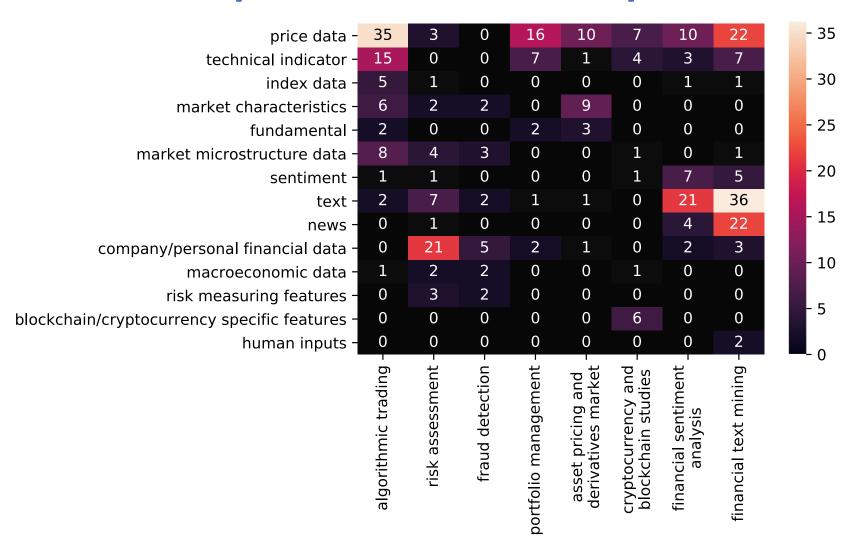
AI, ML, DL



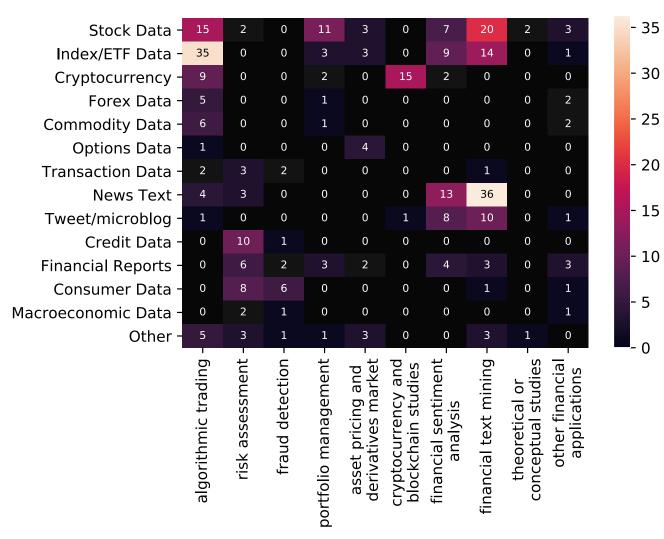
Deep learning for financial applications: Topic-Model Heatmap



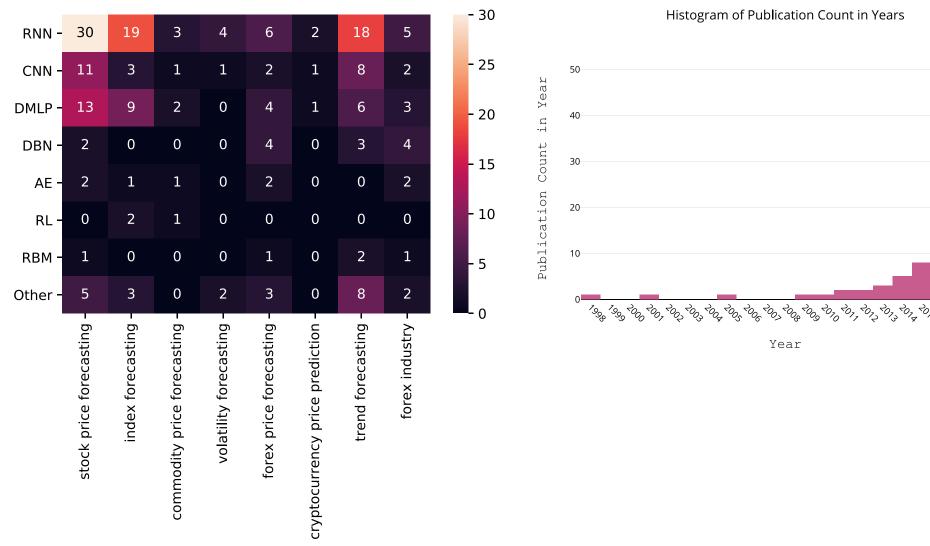
Deep learning for financial applications: Topic-Feature Heatmap



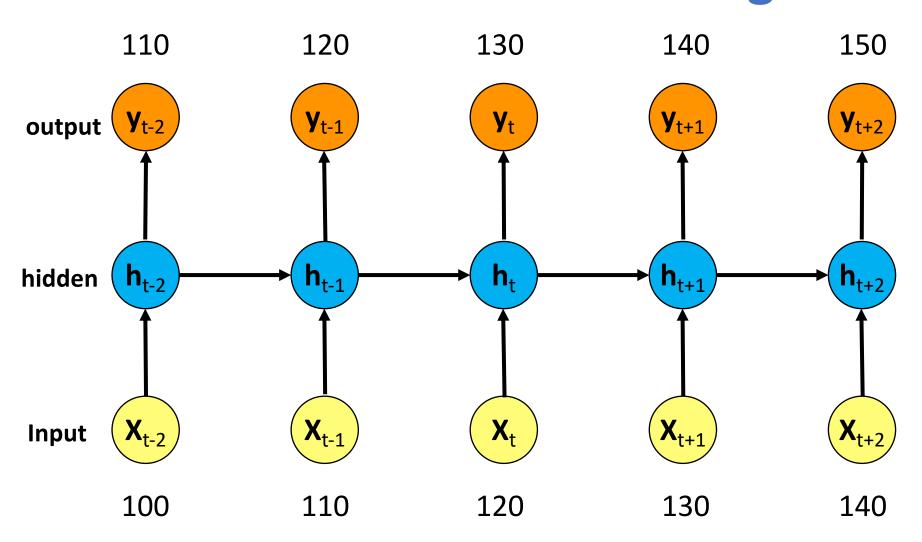
Deep learning for financia applications: Topic-Dataset Heatmap



Financial time series forecasting with deep learning: Topic-model heatmap



Recurrent Neural Networks (RNN) Time Series Forecasting



Deep Learning

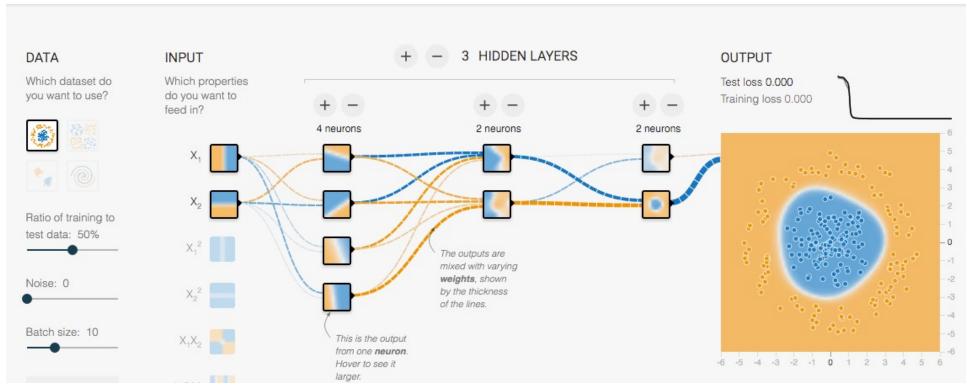
Deep Learning and Neural Networks



TensorFlow Playground

Tinker With a **Neural Network** Right Here in Your Browser. Don't Worry, You Can't Break It. We Promise.







Tensor

- 3
 - # a rank 0 tensor; this is a scalar with shape []
- [1. ,2., 3.]
 - # a rank 1 tensor; this is a vector with shape [3]
- [[1., 2., 3.], [4., 5., 6.]]
 - # a rank 2 tensor; a matrix with shape [2, 3]
- [[[1., 2., 3.]], [[7., 8., 9.]]]
 - # a rank 3 tensor with shape [2, 1, 3]

Scalar

80

Vector

[50 60 70]

Matrix

 50
 60
 70

 55
 65
 75

Tensor

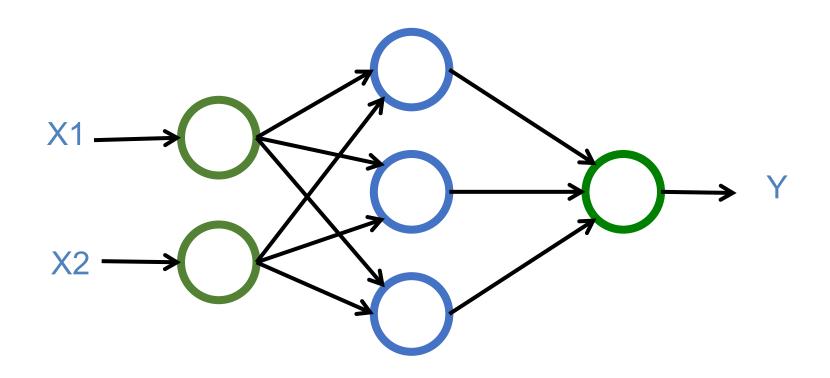
[50 60 70] [70 80 90] [55 65 75] [75 85 95]

Deep Learning and Neural Networks

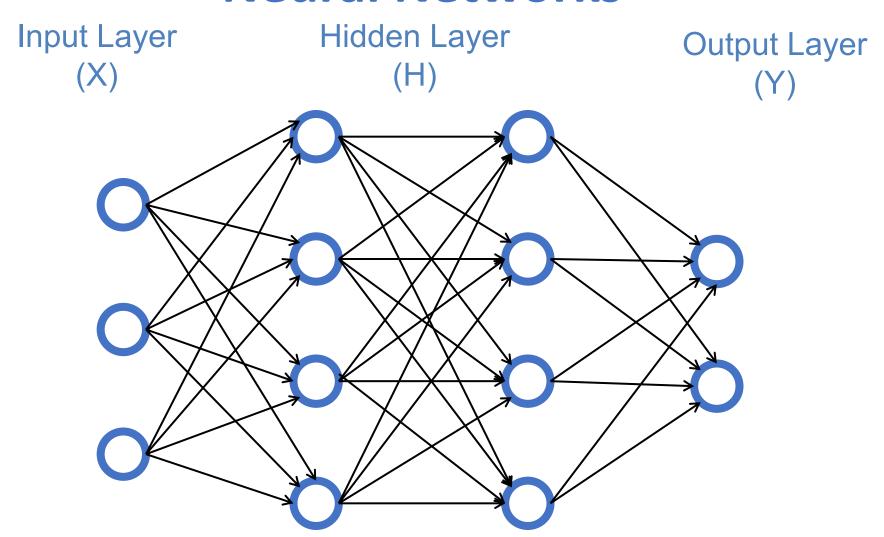
Deep Learning Foundations: Neural Networks

Deep Learning and Neural Networks

Input Layer Hidden Layer Output Layer (X) (H) (Y)



Deep Learning and Neural Networks



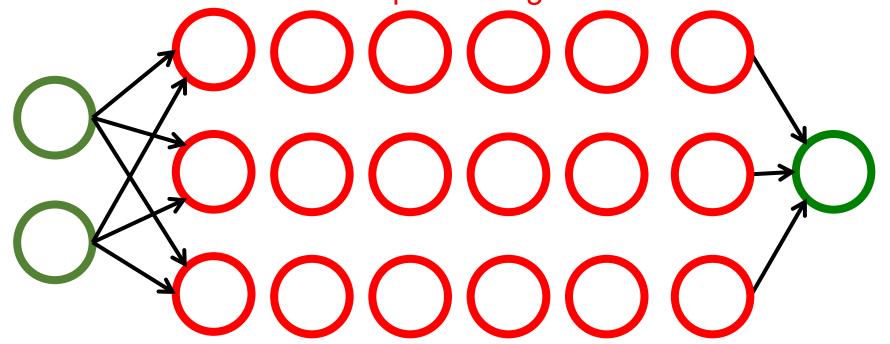
Deep Learning and Neural Networks

Input Layer (X)

Hidden Layers (H)

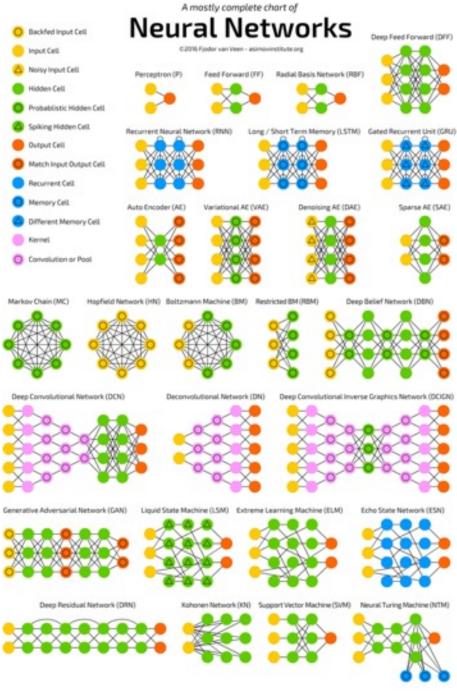
Output Layer (Y)

Deep Neural Networks
Deep Learning



Deep Learning and Deep Neural Networks

Neural Networks (NN)



A mostly complete chart of

Neural Networks

©2016 Fjodor van Veen - asimovinstitute.org

Deep Feed Forward (DFF)







- Hidden Cell
- Probablistic Hidden Cell

Backfed Input Cell

- Spiking Hidden Cell
- Output Cell
- Match Input Output Cell
- Recurrent Cell
- Memory Cell
- Different Memory Cell
- Kernel
- Convolution or Pool











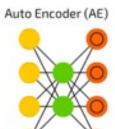


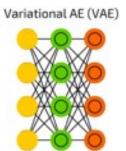


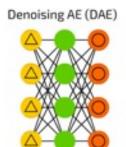


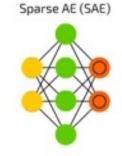






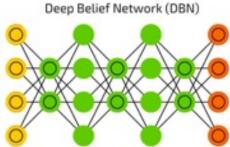






Markov Chain (MC)

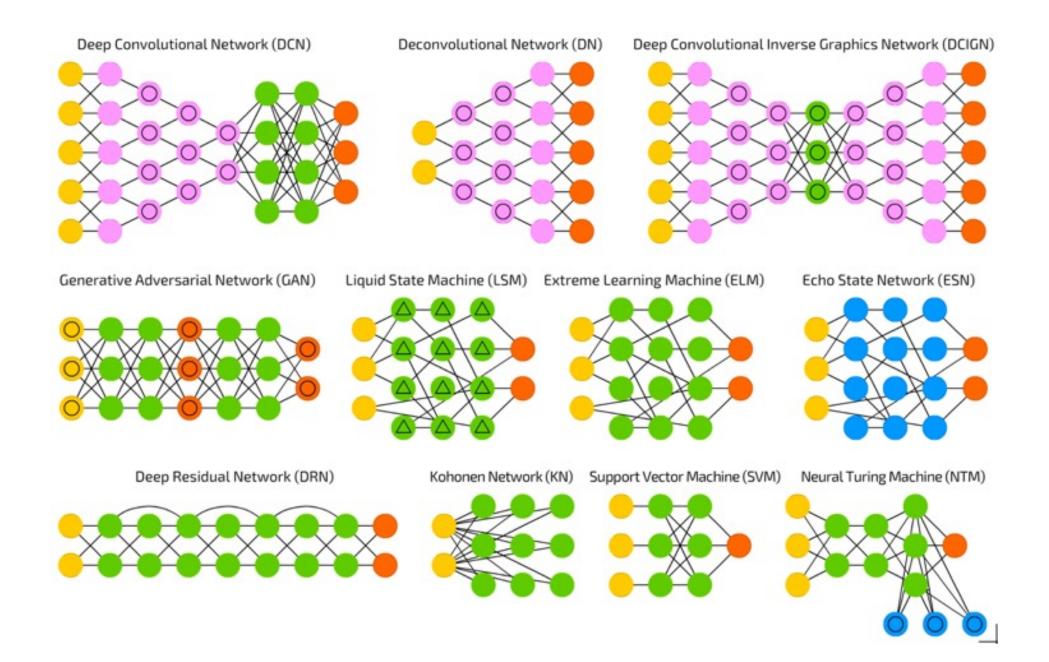
Hopfield Network (HN) Boltzmann Machine (BM) Restricted BM (RBM)





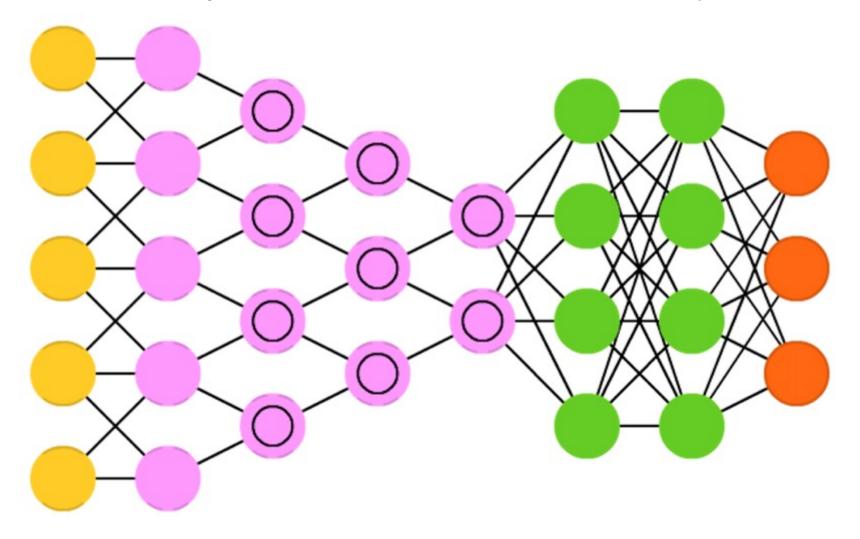




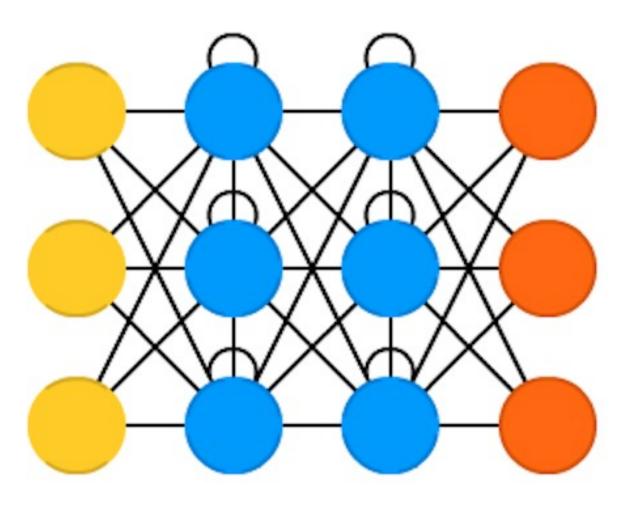


Convolutional Neural Networks (CNN

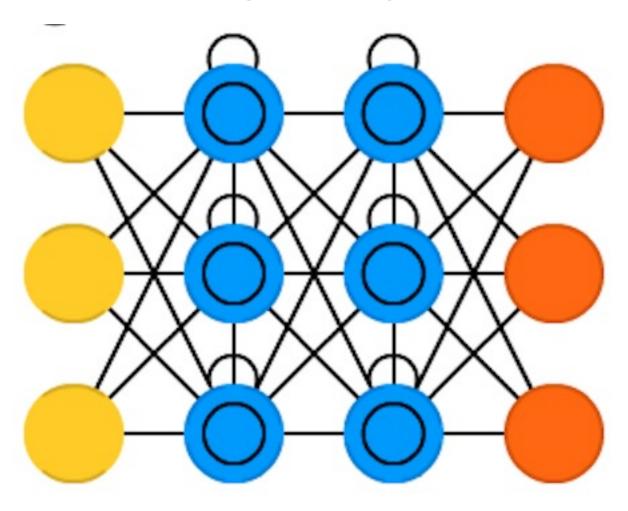
or Deep Convolutional Neural Networks, DCNN)



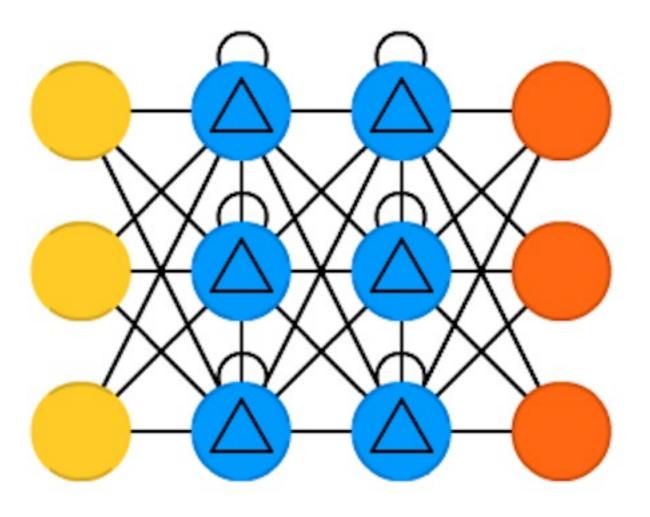
Recurrent Neural Networks (RNN)



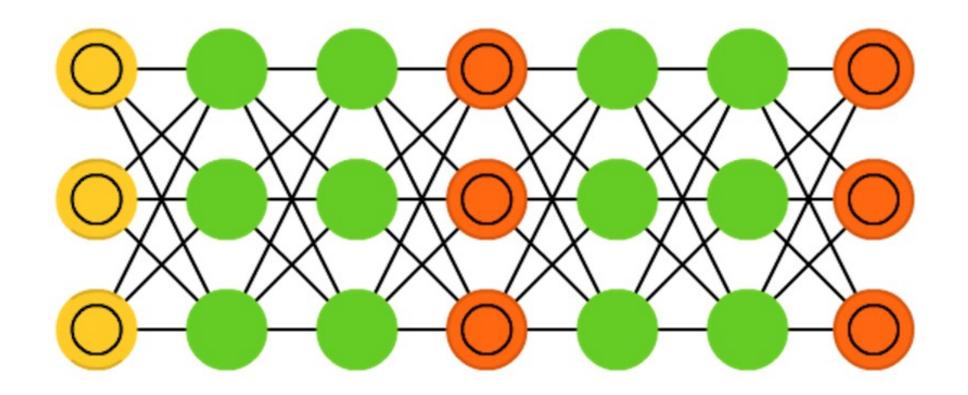
Long / Short Term Memory (LSTM)



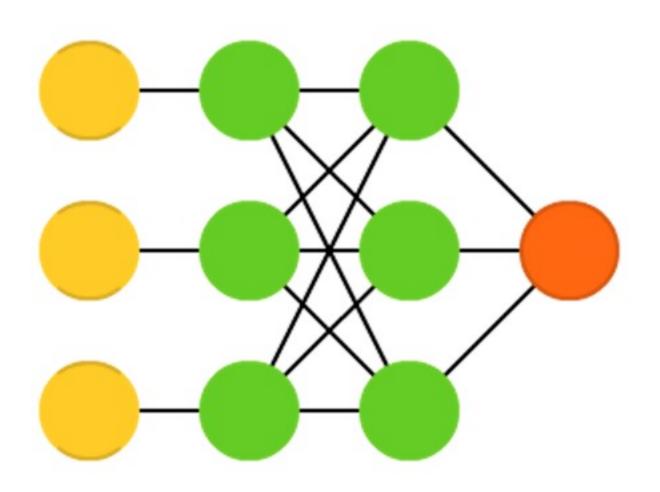
Gated Recurrent Units (GRU)



Generative Adversarial Networks (GAN)



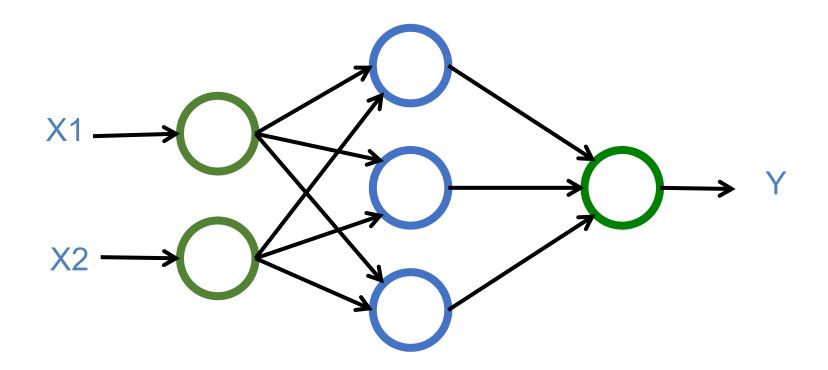
Support Vector Machines (SVM)



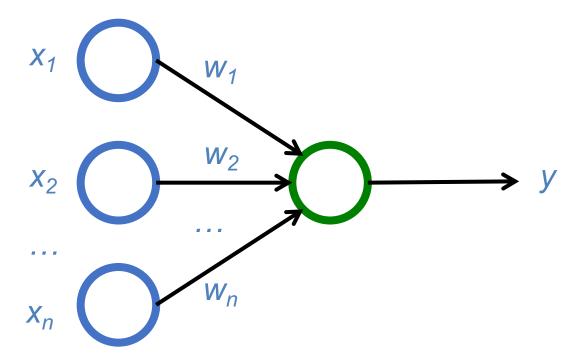
Neural Networks

(X)

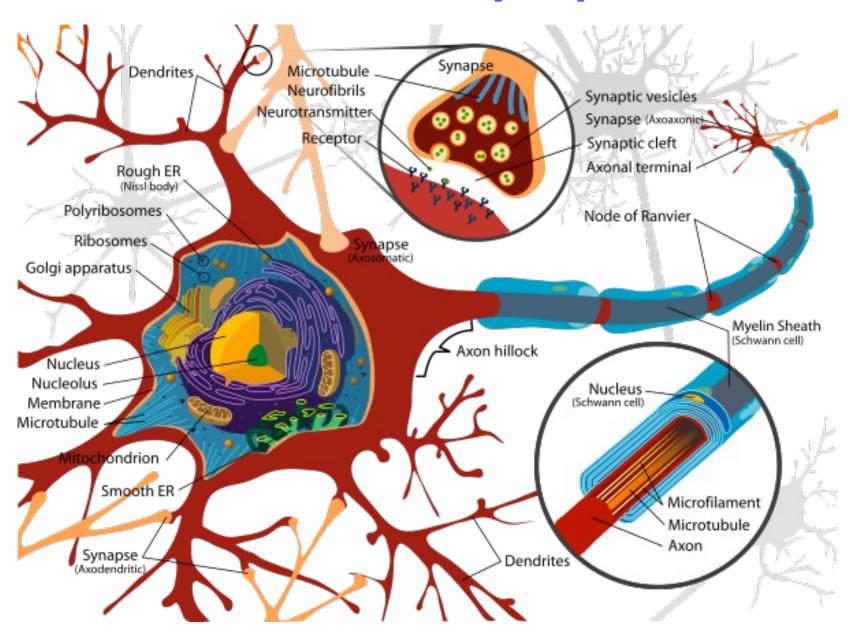
Input Layer Hidden Layer Output Layer



The Neuron



Neuron and Synapse



The Neuron

$$y = F\left(\sum_{i} w_{i} x_{i}\right)$$

$$x_{1}$$

$$x_{2}$$

$$w_{2}$$

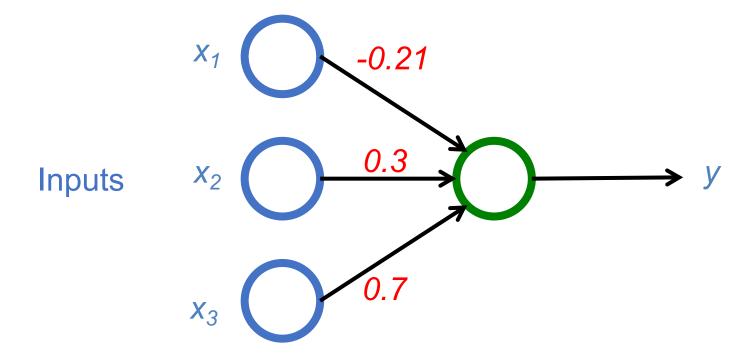
$$x_{n}$$

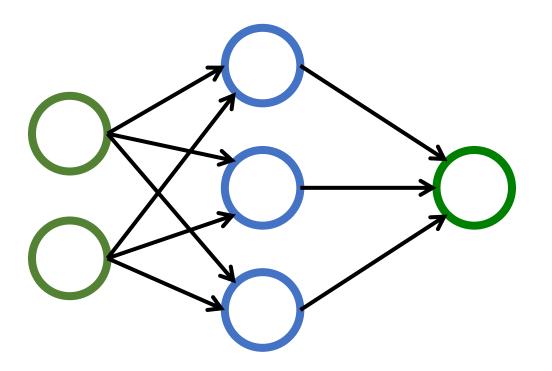
$$w_{n}$$

$$F(x) = \max(0, x)$$

$$y = max(0, -0.21 * x_1 + 0.3 * x_2 + 0.7 * x_3)$$

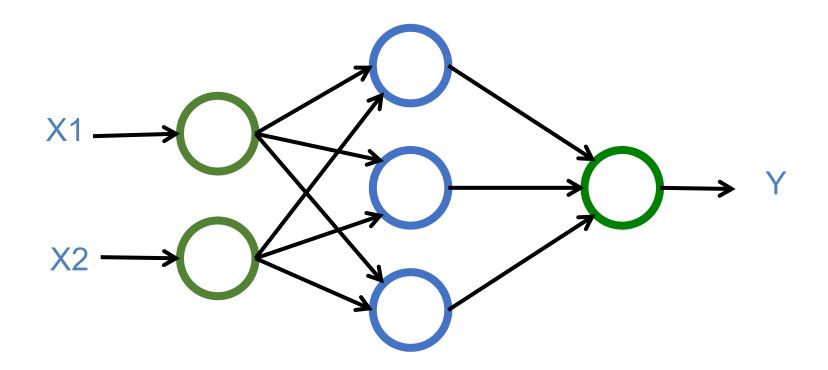
Weights





(X)

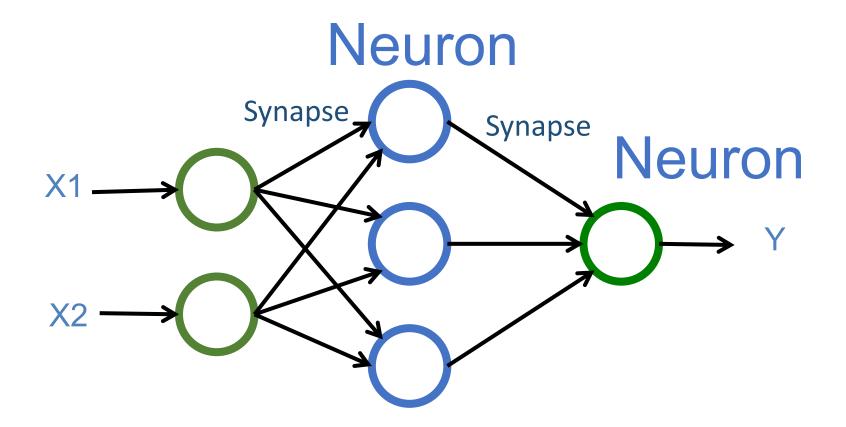
Input Layer Hidden Layer Output Layer

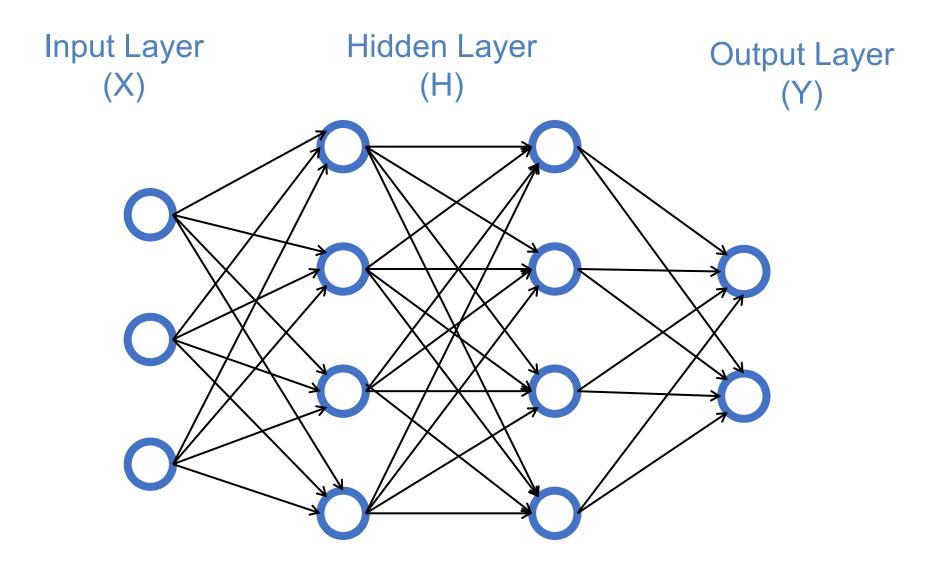


Input Layer Output Layer Hidden Layers (X) **Deep Neural Networks Deep Learning**

Input Layer (X)

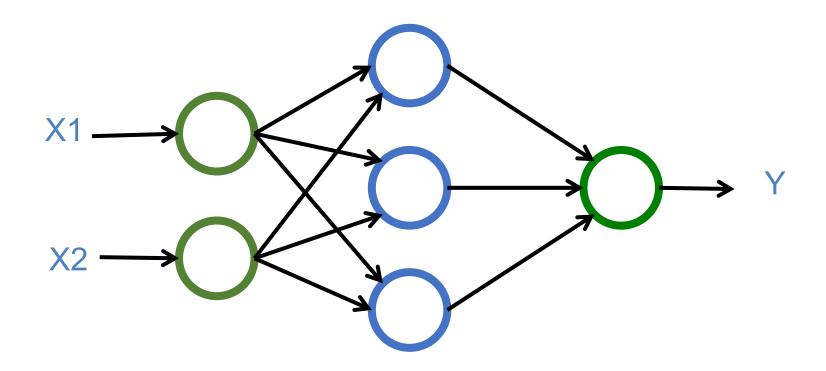
Hidden Layer Output Layer





(X)

Input Layer Hidden Layer Output Layer



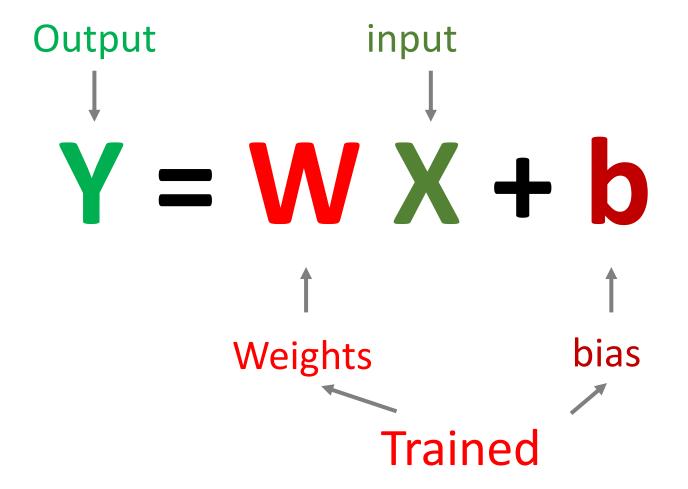
Linear function

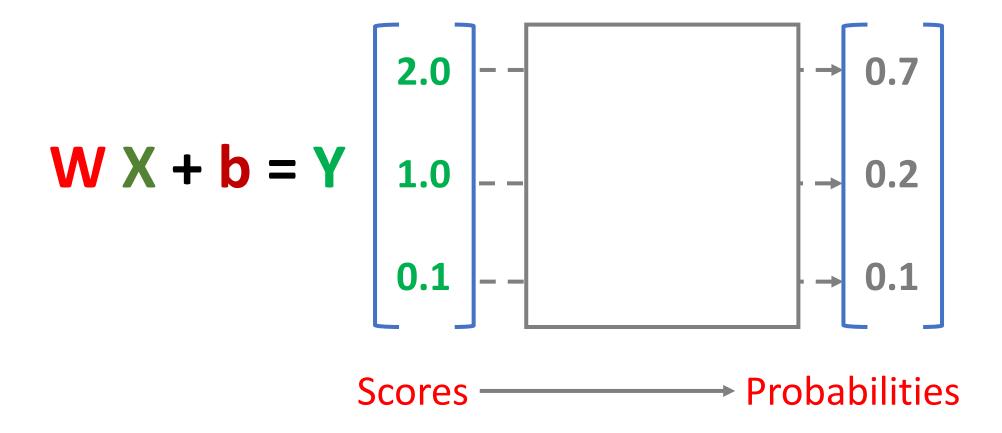
$$y = f(x)$$

$$y = w_1 x + w_0$$

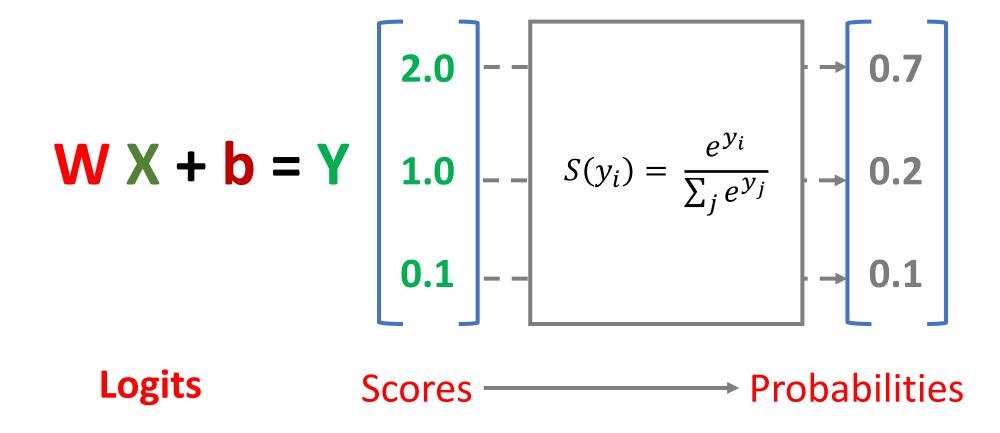
$$h_w(x) = w_1 x + w_0$$

Y = W X + b





SoftMAX



$$S(y_i) = \frac{e^{y_i}}{\sum_j e^{y_i}} = \frac{e^{2.0}}{e^{2.0} + e^{1.0} + e^{0.1}} = \frac{2.7182^{2.0}}{2.7182^{2.0} + 2.7182^{1.0} + 2.7182^{0.1}} = 0.7$$

$$S(y_i) = \frac{e^{y_i}}{\sum_i e^{y_j}} = \frac{e^{1.0}}{e^{2.0} + e^{1.0} + e^{0.1}} = \frac{2.7182^{1.0}}{2.7182^{2.0} + 2.7182^{1.0} + 2.7182^{0.1}} = 0.2$$

$$S(y_i) = \frac{e^{y_i}}{\sum_j e^{y_j}} = \frac{e^{1.0}}{e^{2.0} + e^{1.0} + e^{0.1}} = \frac{2.7182^{1.0}}{2.7182^{2.0} + 2.7182^{1.0} + 2.7182^{0.1}} = 0.2$$

$$S(y_i) = \frac{e^{y_i}}{\sum_j e^{y_j}} = \frac{e^{0.1}}{e^{2.0} + e^{1.0} + e^{0.1}} = \frac{2.7182^{0.1}}{2.7182^{2.0} + 2.7182^{1.0} + 2.7182^{0.1}} = 0.1$$

$$\mathbf{W} \ \mathbf{X} + \mathbf{b} = \mathbf{Y} \begin{bmatrix} 2.0 \\ 1.0 \\ -- \end{bmatrix} S(y_i) = \frac{e^{y_i}}{\sum_j e^{y_j}} \rightarrow \mathbf{0.2}$$

Logits

Probabilities Scores

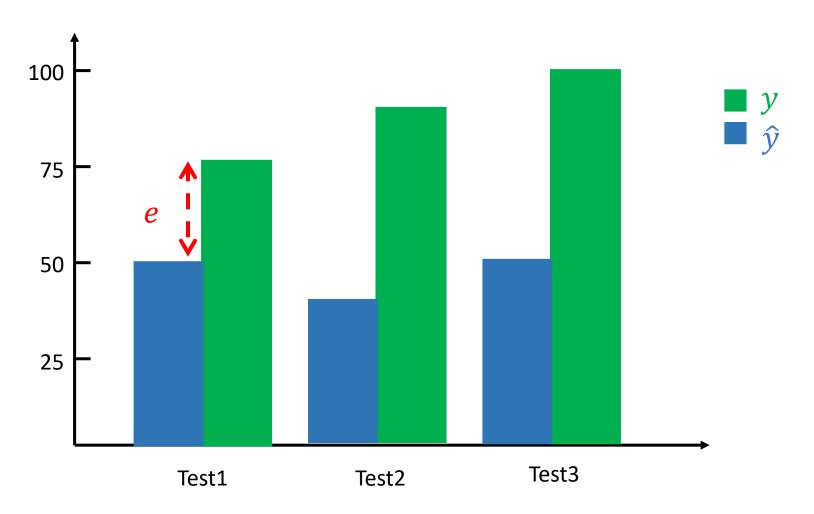
Training a Network =

Minimize the Cost Function

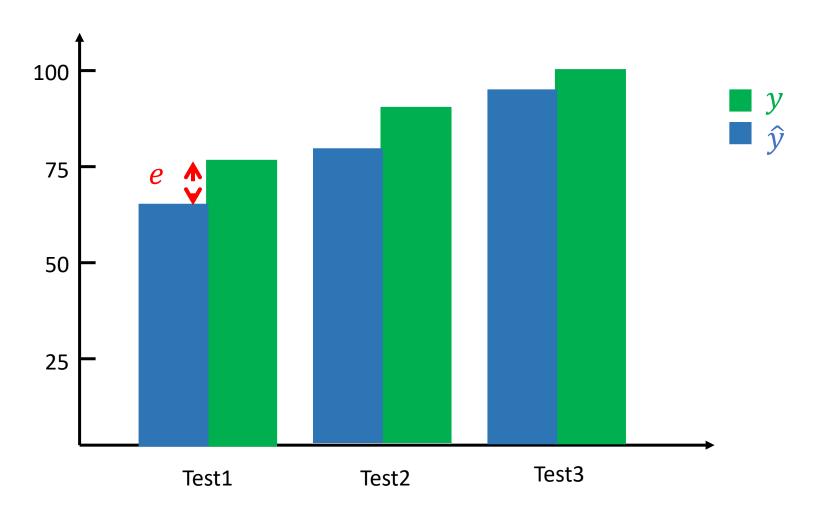
Training a Network

Minimize the Cost Function Minimize the Loss Function

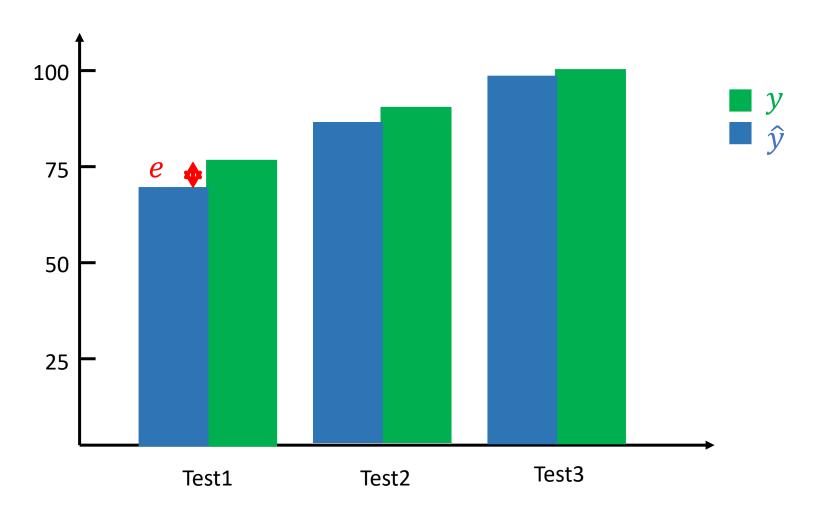
Error = Predict Y - Actual Y Error : Cost : Loss



Error = Predict Y - Actual Y Error : Cost : Loss

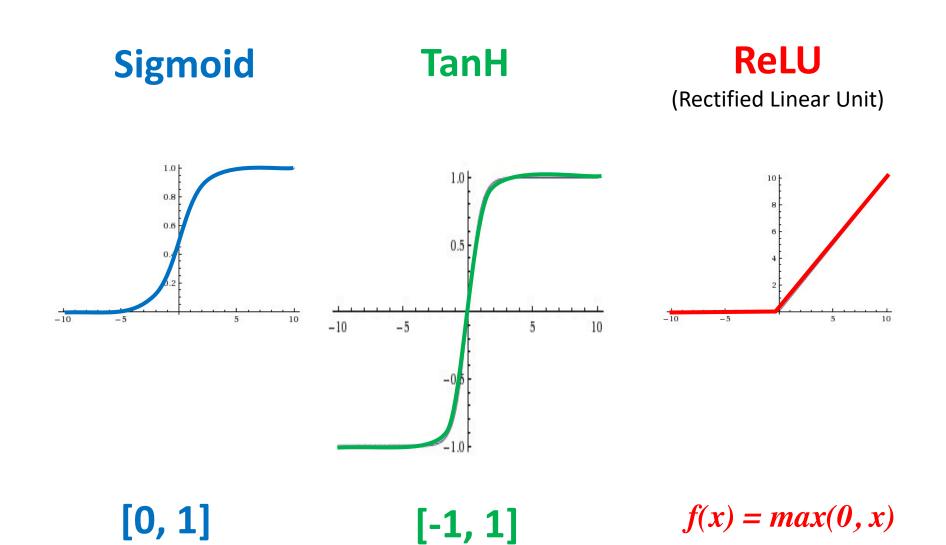


Error = Predict Y - Actual Y Error : Cost : Loss

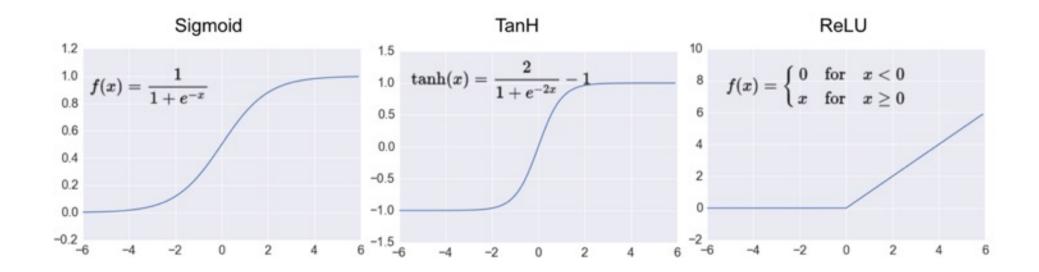


Activation Functions

Activation Functions



Activation Functions



Loss Function

Binary Classification: 2 Class

Activation Function: Sigmoid

Loss Function: Binary Cross-Entropy

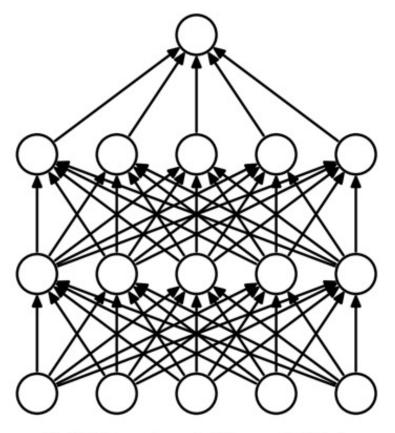
Multiple Classification: 10 Class

Activation Function: SoftMAX

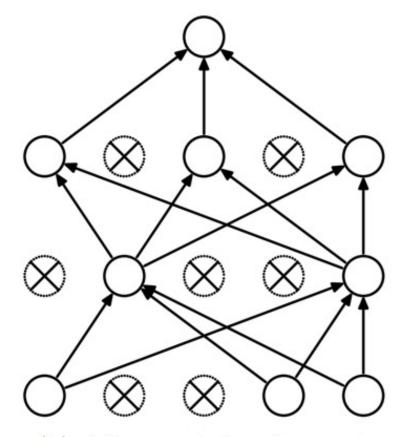
Loss Function:
Categorical Cross-Entropy

Dropout

Dropout: a simple way to prevent neural networks from overfitting



(a) Standard Neural Net



(b) After applying dropout.

Learning Algorithm

While not done:

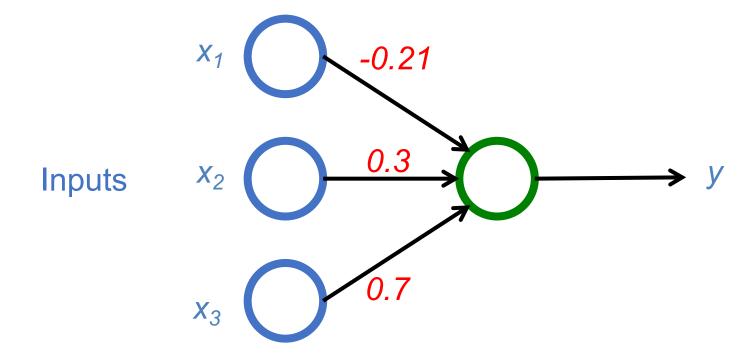
Pick a random training example "(input, label)"

Run neural network on "input"

Adjust weights on edges to make output closer to "label"

$$y = max(0, -0.21 * x_1 + 0.3 * x_2 + 0.7 * x_3)$$

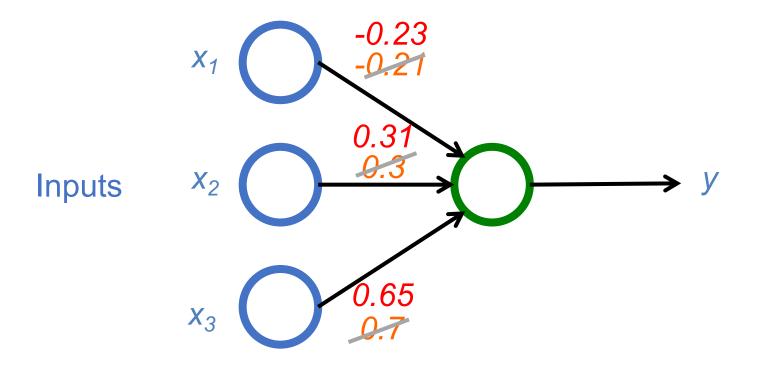
Weights



Next time:

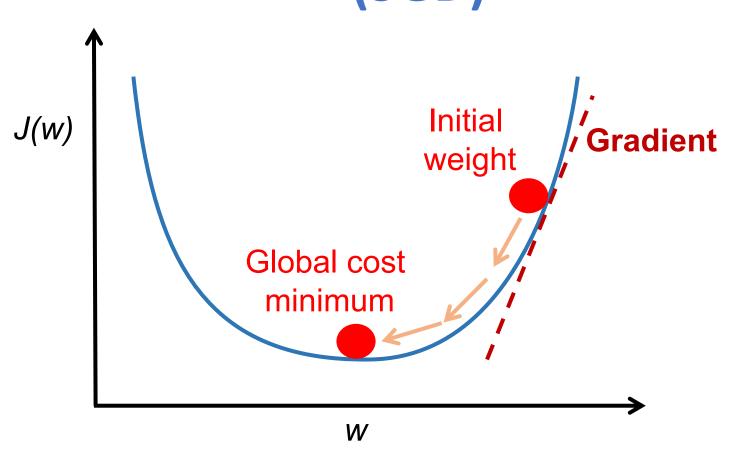
$$y = max (0, -0.23 * x_1 + 0.31 * x_2 + 0.65 * x_3)$$

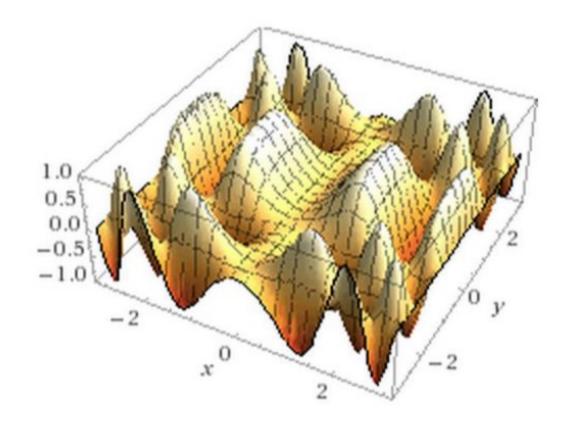
 $y = max (0, -0.21 * x_1 + 0.3 * x_2 + 0.7 * x_3)$
Weights



Optimizer:

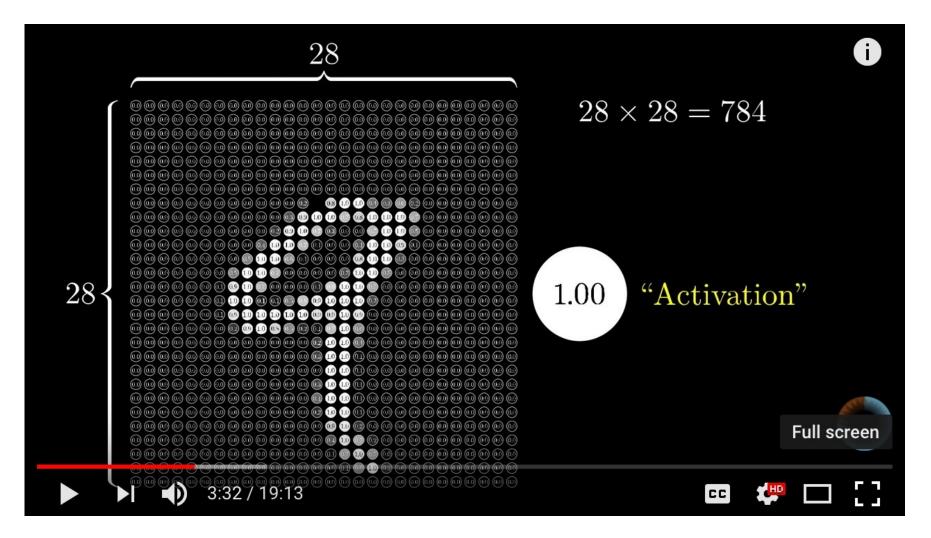
Stochastic Gradient Descent (SGD)



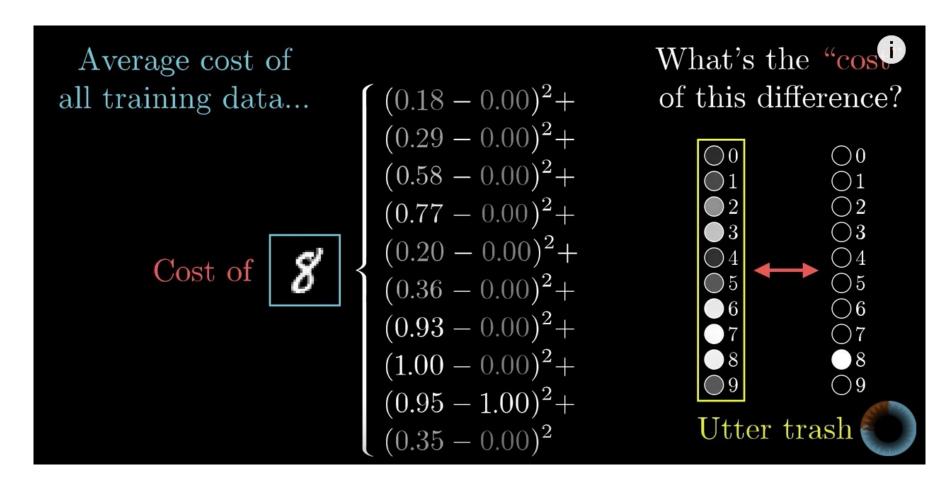


This shows a function of 2 variables: real neural nets are functions of hundreds of millions of variables!

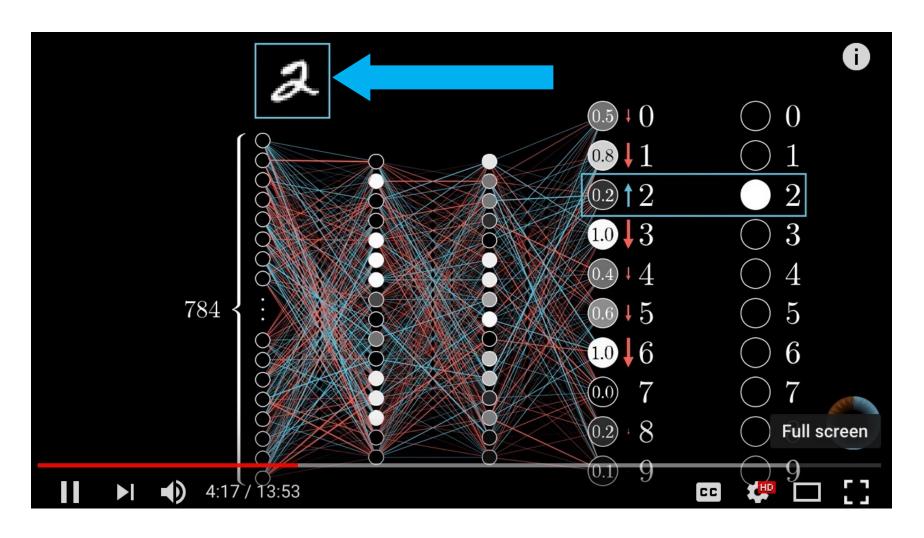
Neural Network and Deep Learning



Gradient Descent how neural networks learn



Backpropagation



Learning Algorithm

While not done:

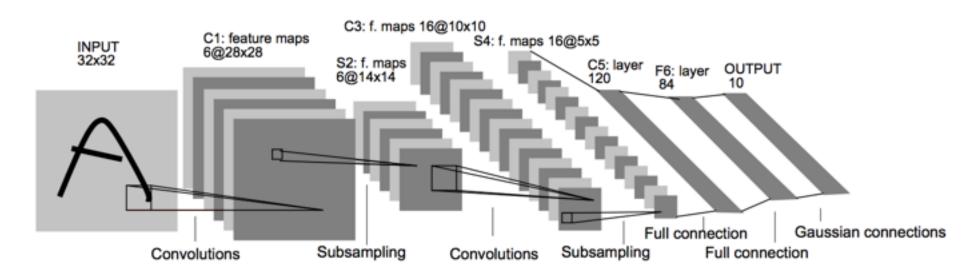
Pick a random training example "(input, label)"

Run neural network on "input"

Adjust weights on edges to make output closer to "label"

Convolutional Neural Networks (CNN)

Convolutional Neural Networks (CNN)

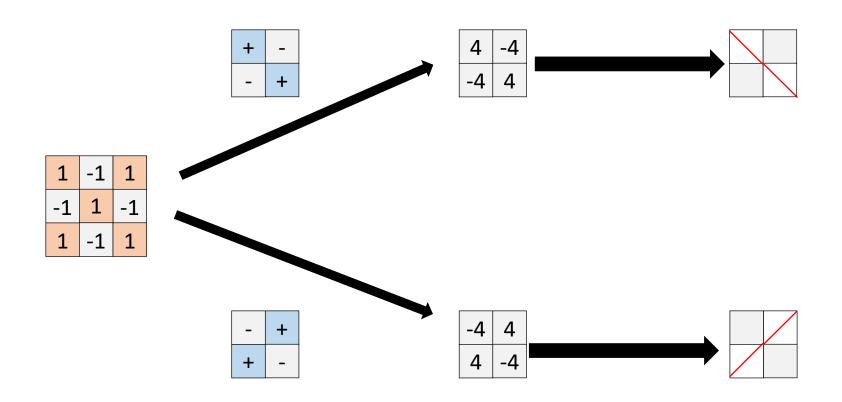


Architecture of LeNet-5 (7 Layers) (LeCun et al., 1998)

Source: http://yann.lecun.com/exdb/publis/pdf/lecun-01a.pdf

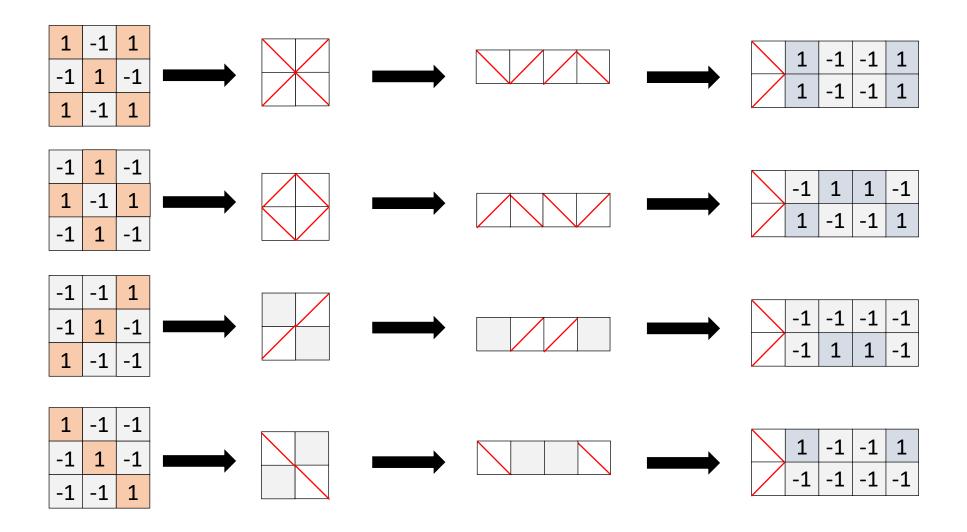
Convolutional Neural Networks (CNN)

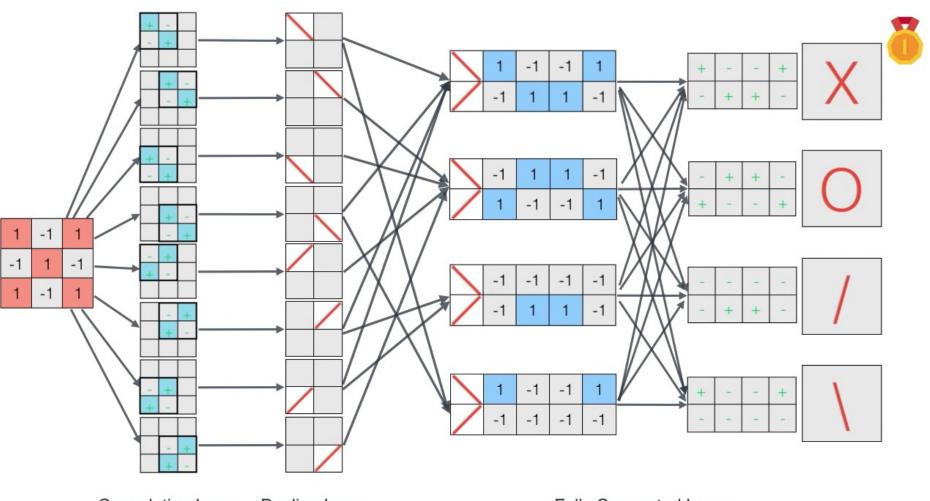
- Convolution
- Pooling
- Fully Connection (FC) (Flattening)



Convolution Layer

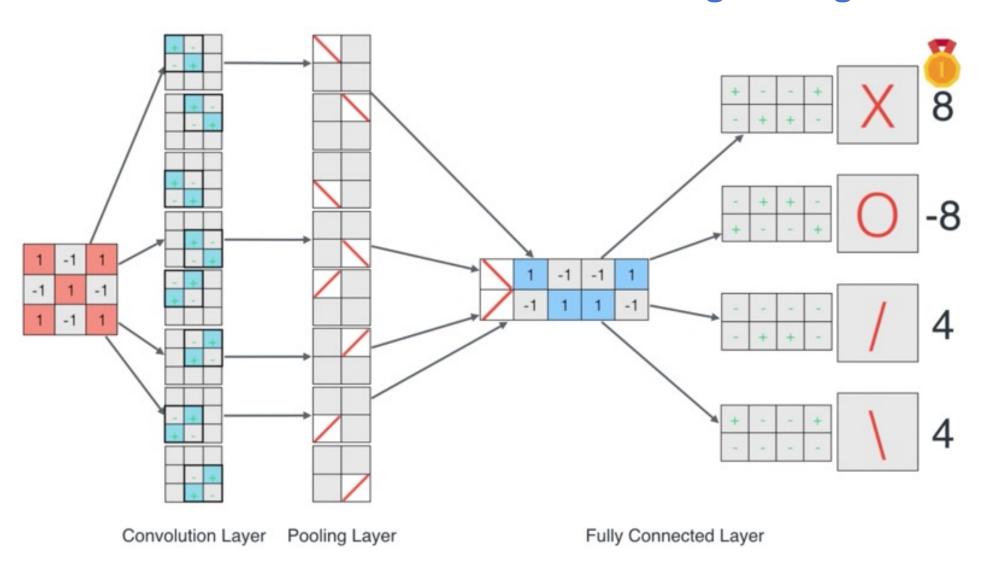
Pooling Layer



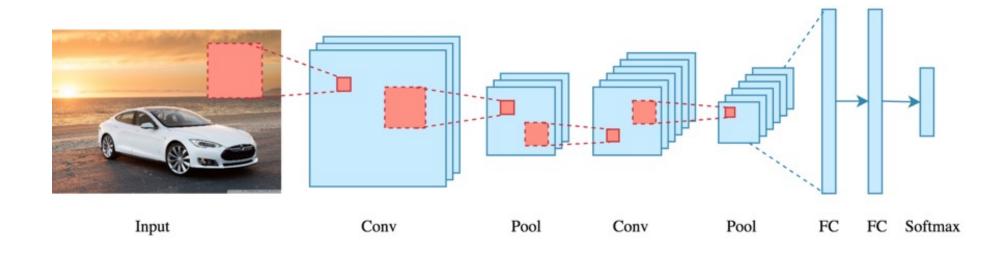


Convolution Layer Pooling Layer

Fully Connected Layer



CNN Architecture



Convolution is a mathematical operation to merge two sets of information 3x3 convolution

1	1	1	0	0
0	1	1	1	0
0	О	1	1	1
0	О	1	1	0
0	1	1	0	0

1	0	1
0	1	0
1	0	1

Input

Filter / Kernel

CNN Convolution Layer Input x Filter --> Feature Map

receptive field: 3x3

1x1	1x0	1x1	0	О
0x0	1x1	1x0	1	О
0x1	0x0	1x1	1	1
0	0	1	1	О
0	1	1	0	0

4	

Input x Filter

CNN Convolution Layer Input x Filter --> Feature Map

receptive field: 3x3

1	1x1	1x0	0x1	0
0	1x0	1x1	1x0	0
0	0x1	1x0	1x1	1
0	О	1	1	0
0	1	1	0	0

4	3	

Input x Filter

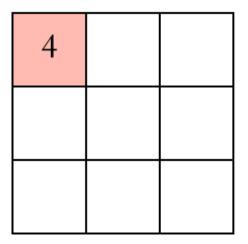




nput

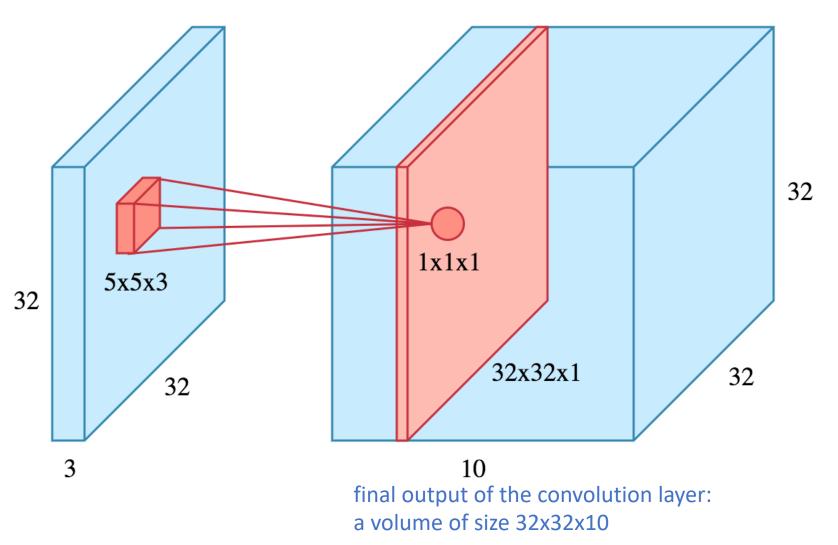
Filter / Kernel

1x1	1x0	1x1	0	0
0x0	1x1	1x0	1	0
0x1	0x0	1x1	1	1
0	0	1	1	0
0	1	1	0	0

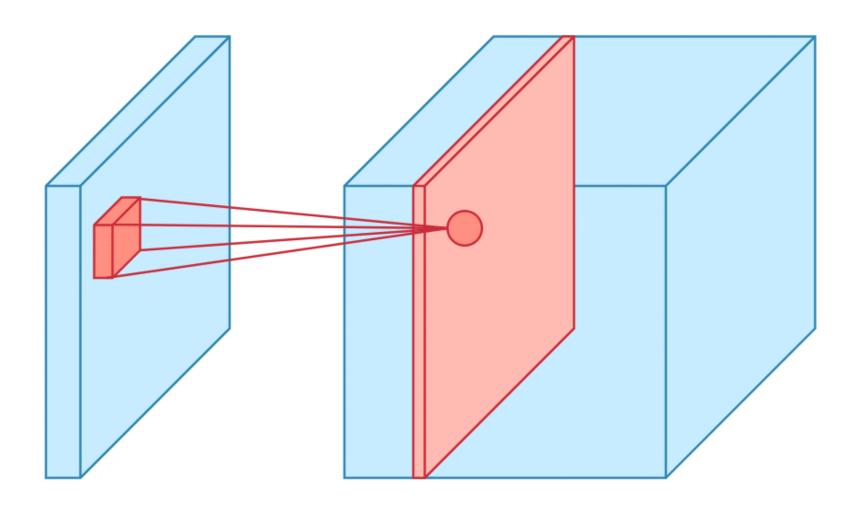


Example convolution operation shown in 2D using a 3x3 filter

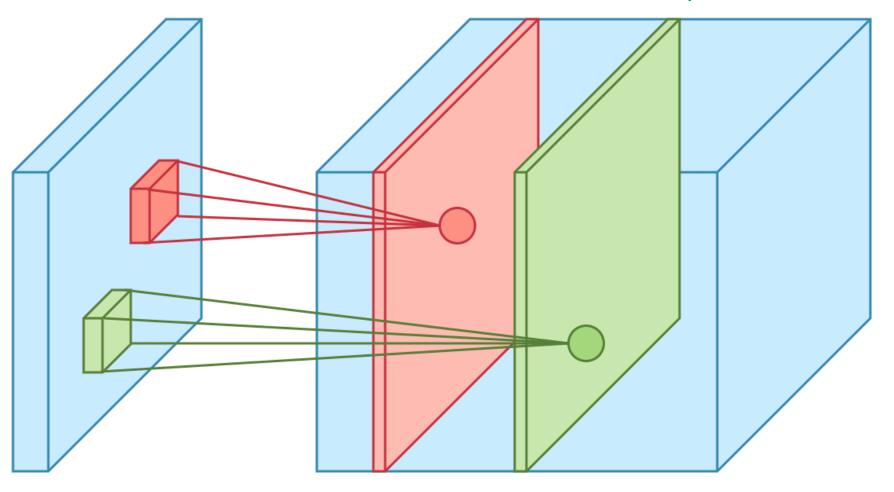
10 different filters 10 feature maps of size 32x32x1



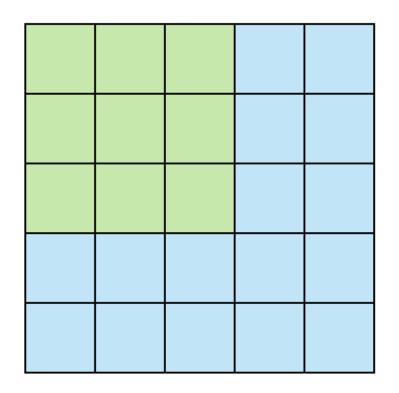
CNN Convolution Layer Sliding operation at 4 locations

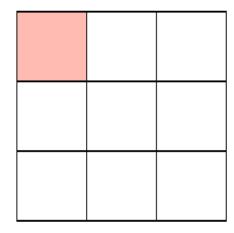


two feature maps



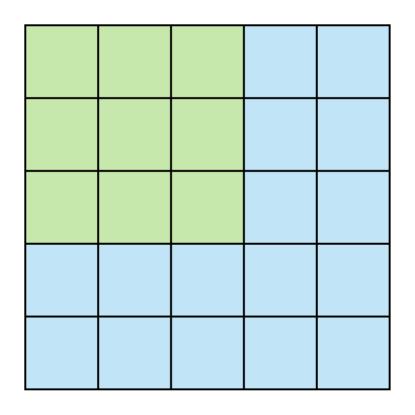
Stride specifies how much we move the convolution filter at each step

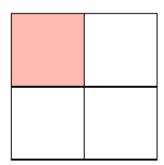




Stride 1

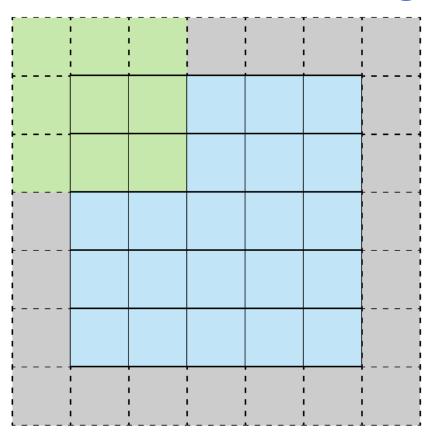
Stride specifies how much we move the convolution filter at each step

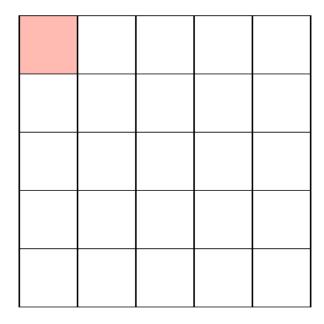




Stride 2

Stride 1 with **Padding**

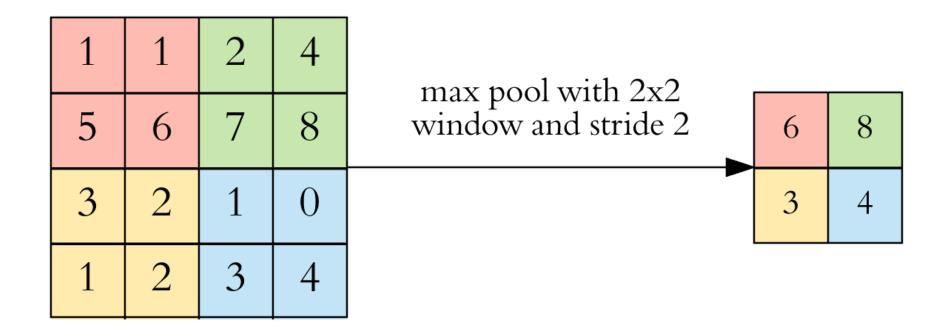




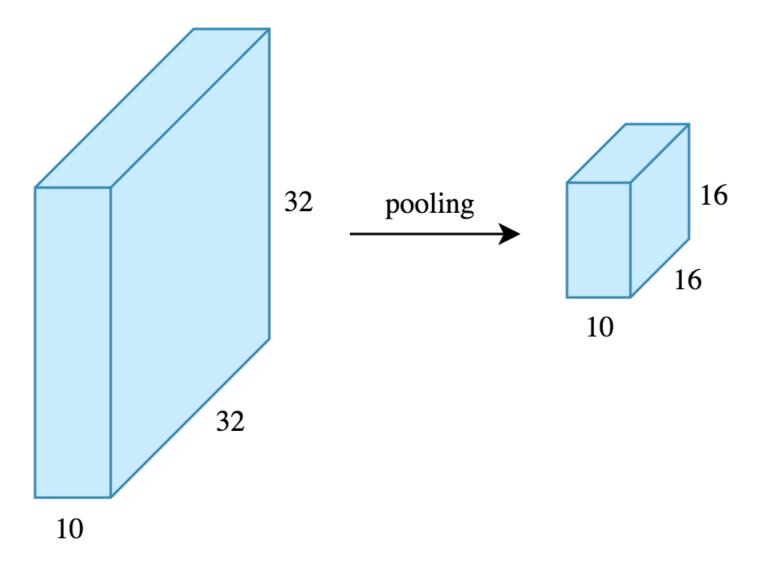
Stride 1 with Padding

CNN Pooling Layer

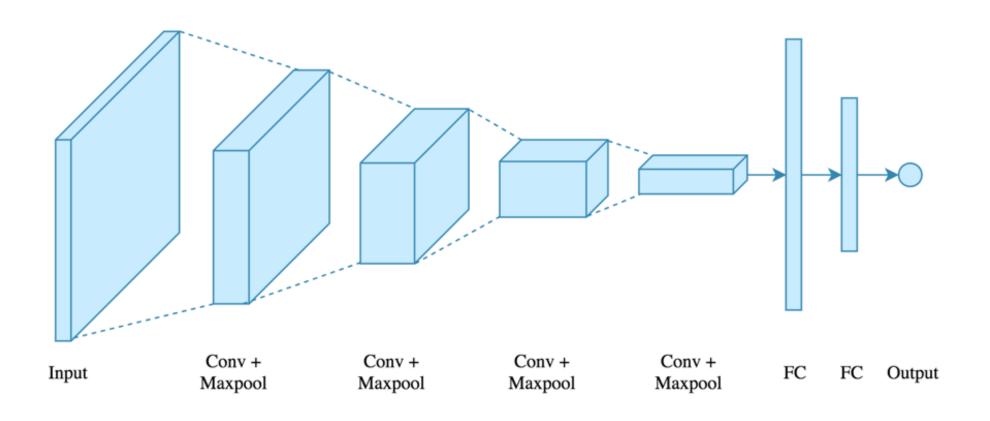
Max Pooling



CNN Pooling Layer



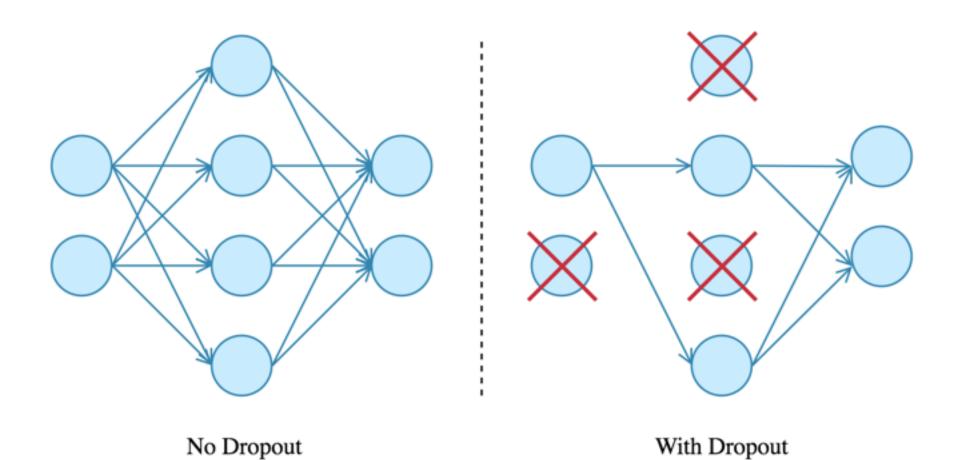
CNN Architecture 4 convolution + pooling layers, followed by 2 fully connected layers



CNN Architecture 4 convolution + pooling layers, followed by 2 fully connected layers

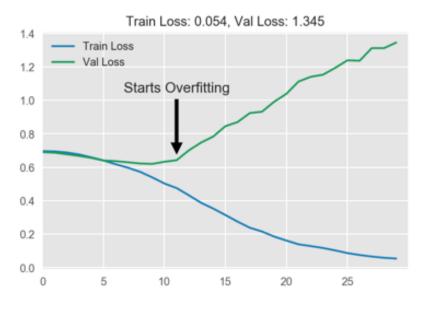
https://gist.github.com/ardendertat/0fc5515057c47e7386fe04e9334504e3

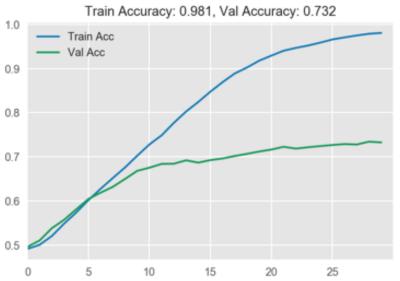
Dropout



Source: Arden Dertat (2017), Applied Deep Learning - Part 4: Convolutional Neural Networks, https://towardsdatascience.com/applied-deep-learning-part-4-convolutional-neural-networks-584bc134c1e2

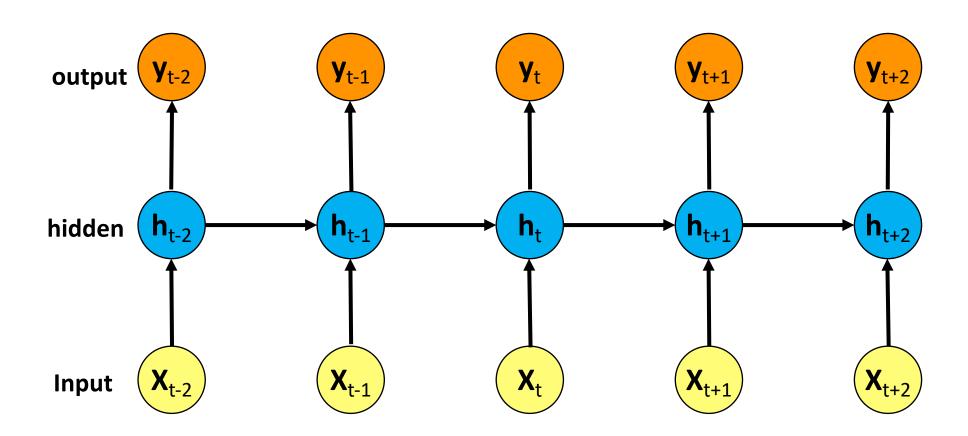
Model Performance



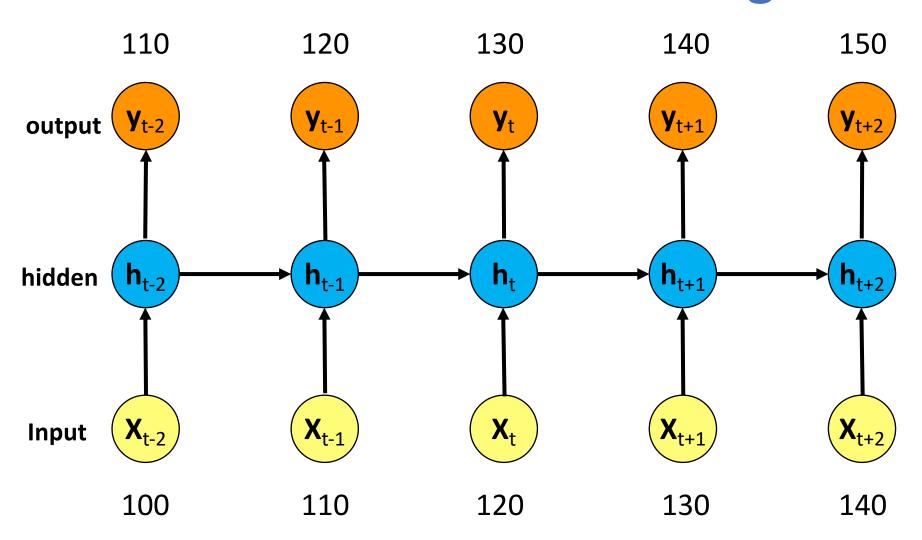


Recurrent Neural Networks (RNN)

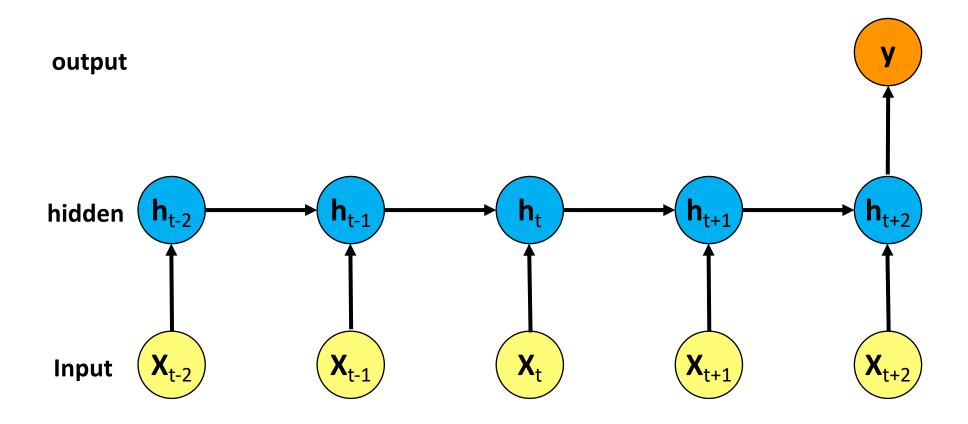
Recurrent Neural Networks (RNN)



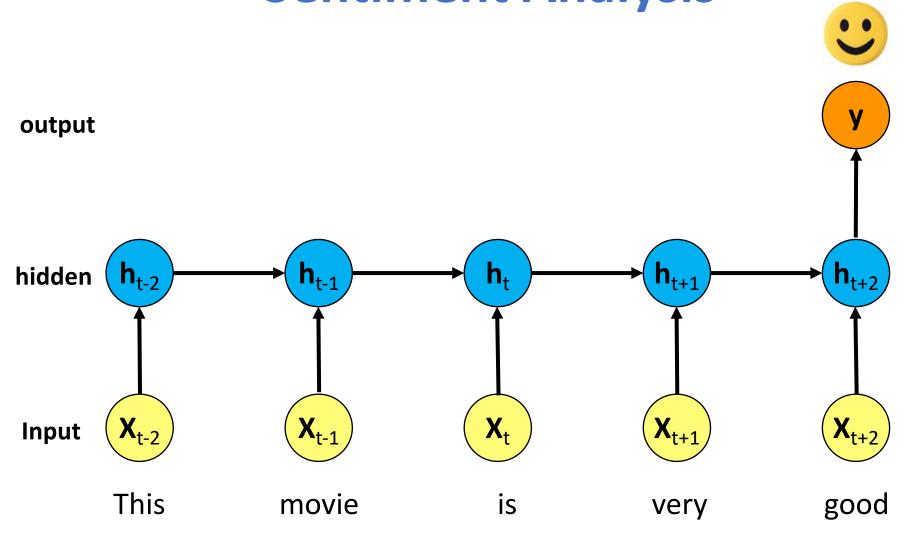
Recurrent Neural Networks (RNN) Time Series Forecasting



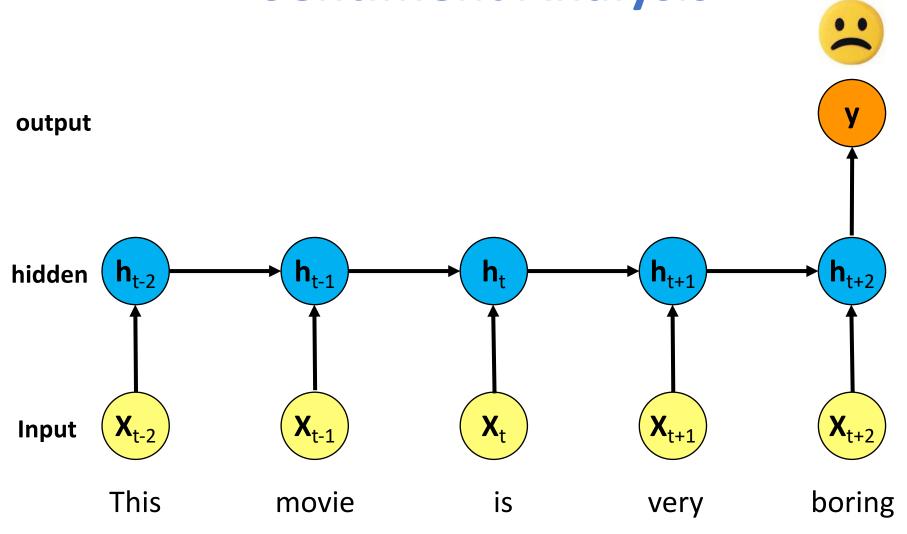
Recurrent Neural Networks (RNN)



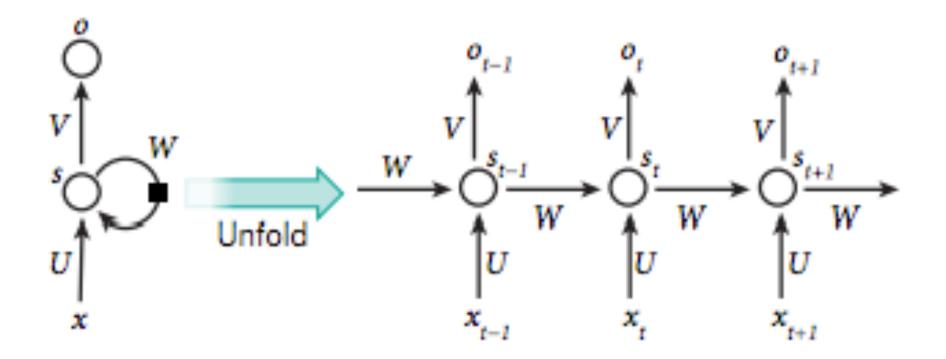
Recurrent Neural Networks (RNN) Sentiment Analysis



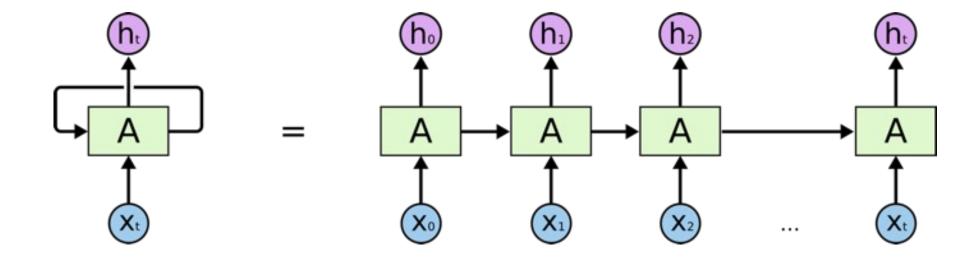
Recurrent Neural Networks (RNN) Sentiment Analysis



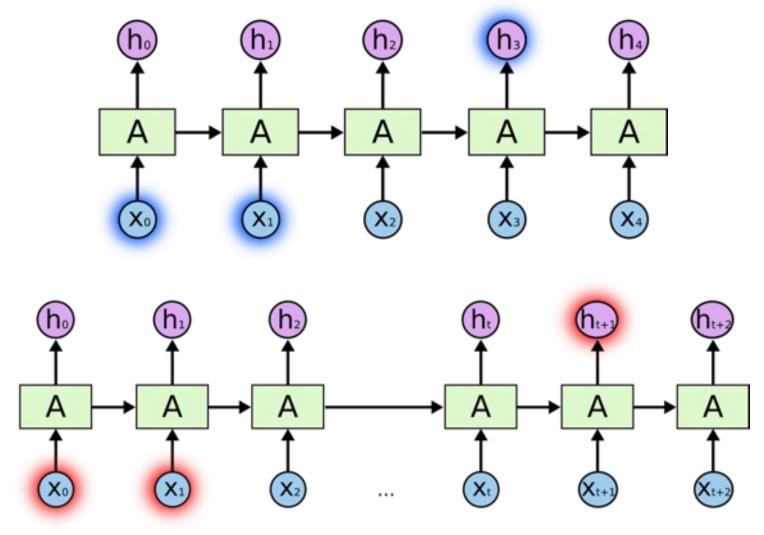
Recurrent Neural Network (RNN)



RNN

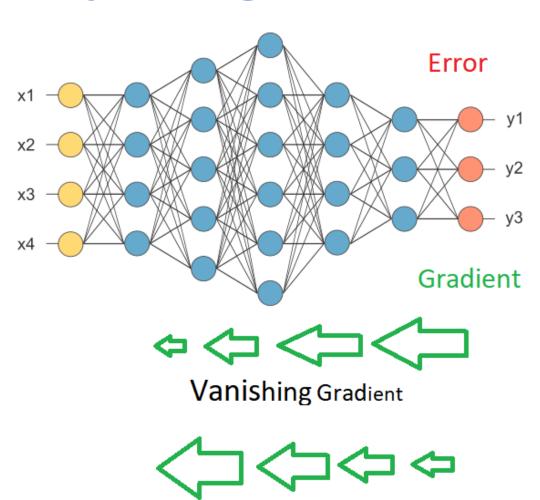


RNN long-term dependencies

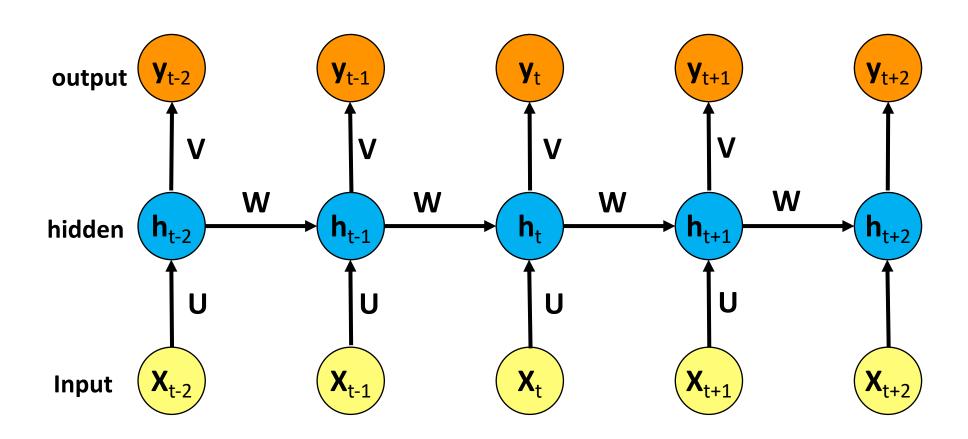


I grew up in France... I speak fluent French.

Vanishing Gradient Exploding Gradient

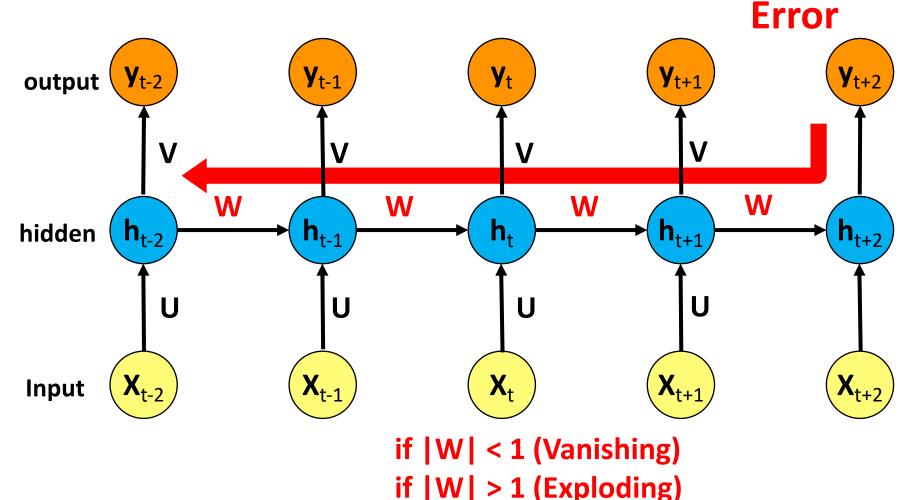


Recurrent Neural Networks (RNN)



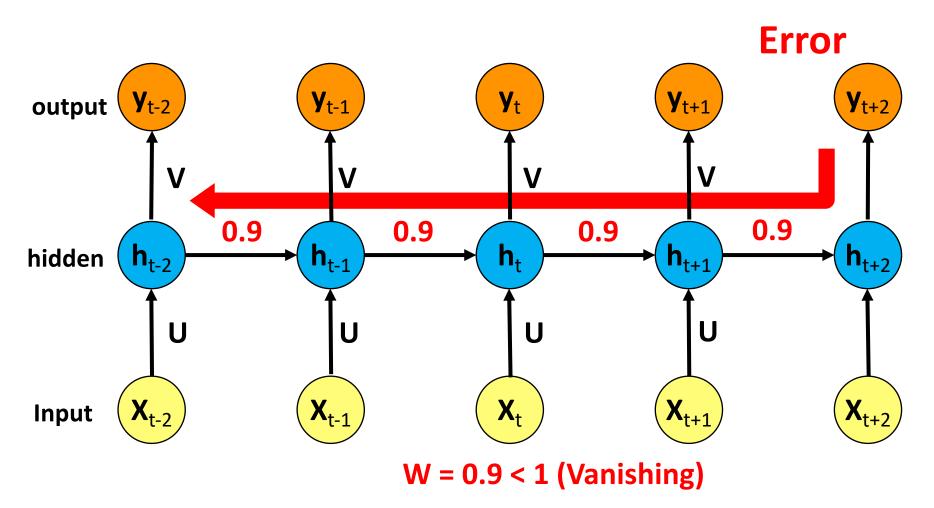
RNN

Vanishing Gradient problem Exploding Gradient problem



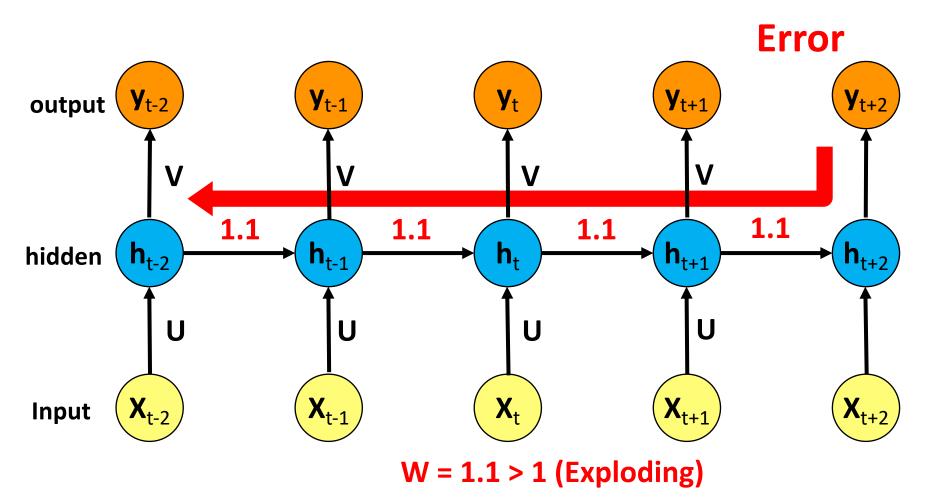
RNN

Vanishing Gradient problem

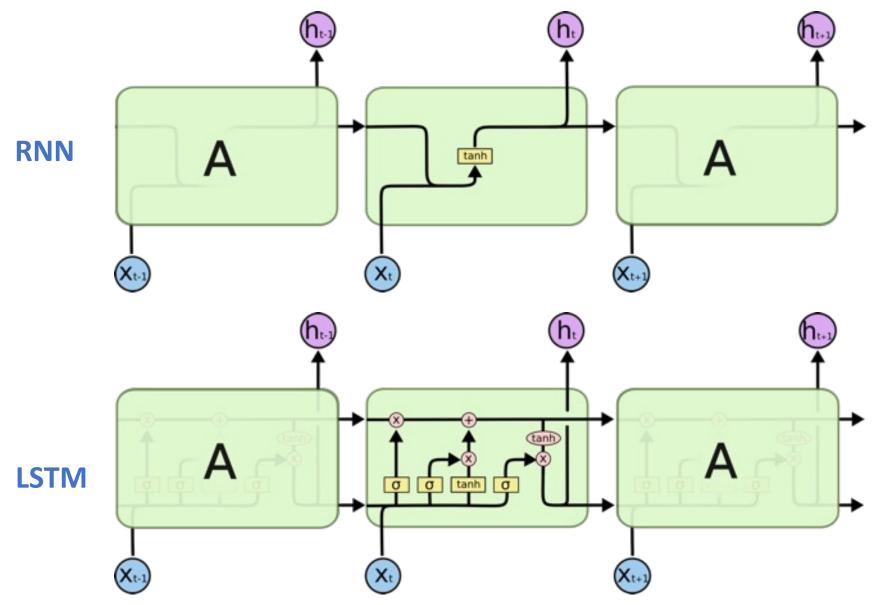


RNN

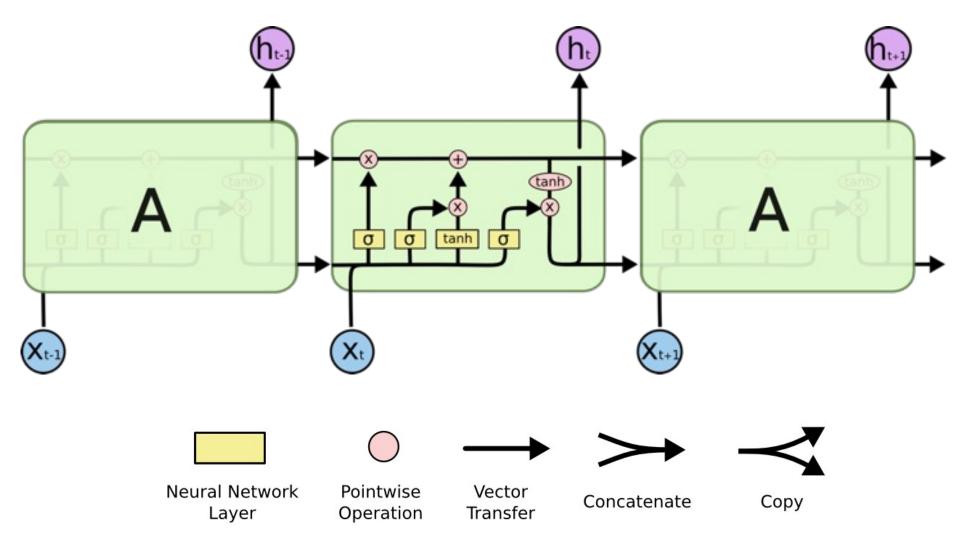
Exploding Gradient problem



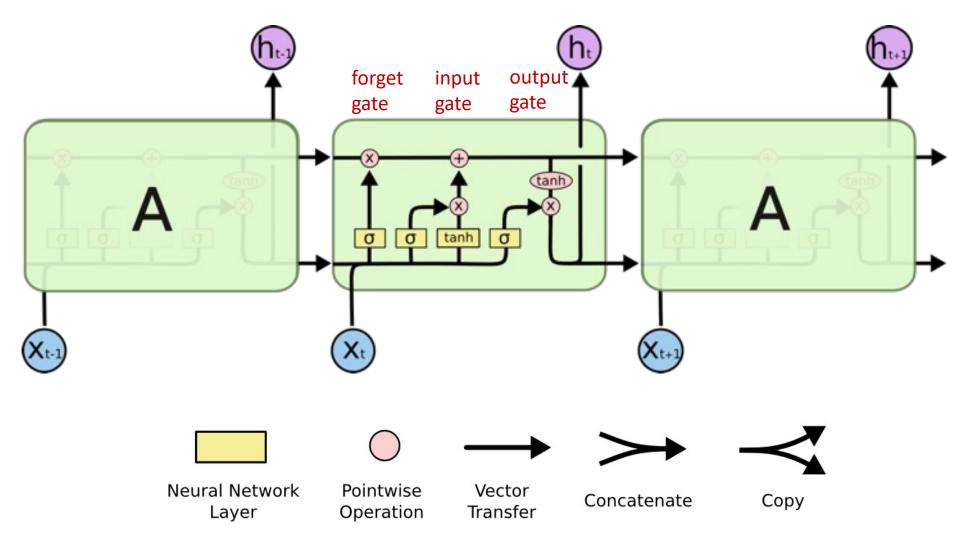
RNN LSTM



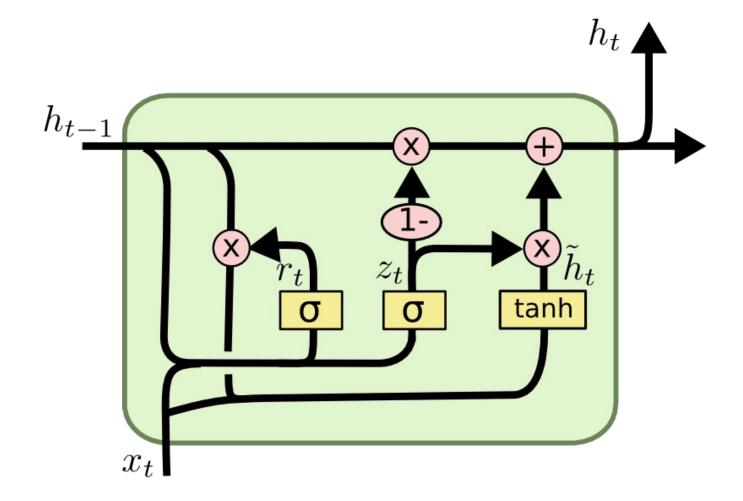
Long Short Term Memory (LSTM)



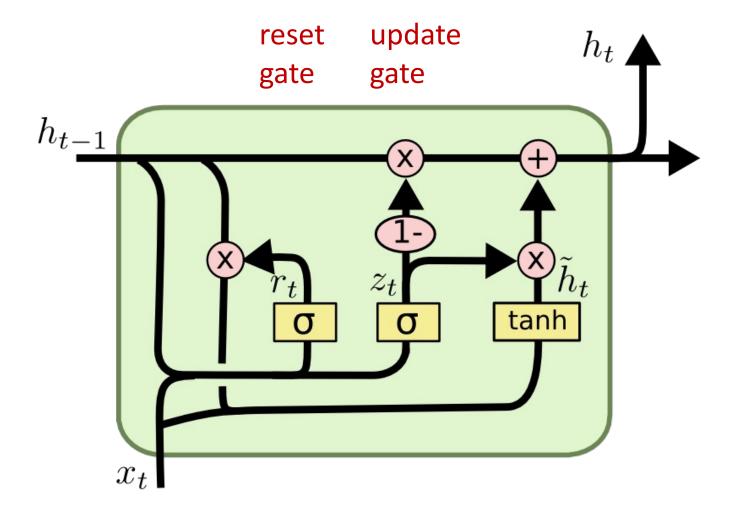
Long Short Term Memory (LSTM)



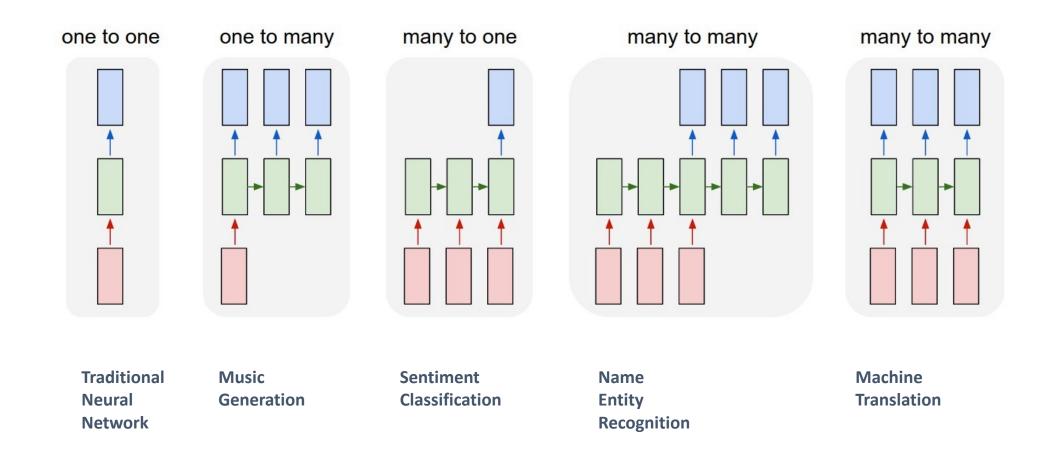
Gated Recurrent Unit (GRU)



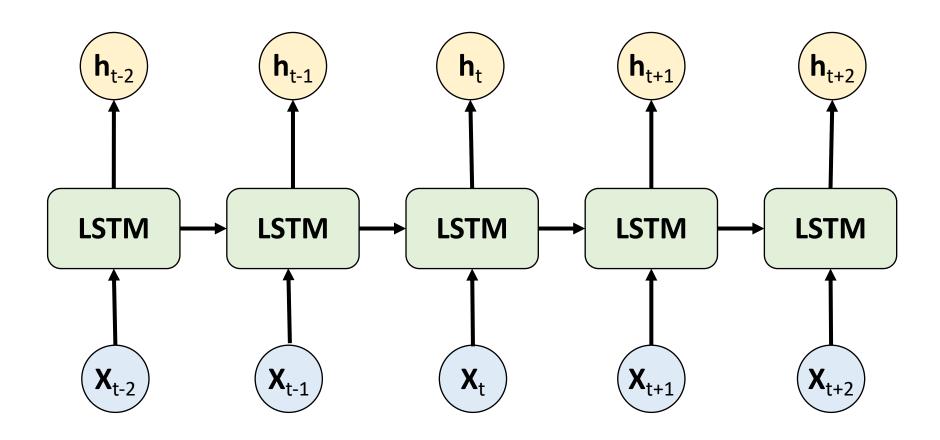
Gated Recurrent Unit (GRU)



LSTM Recurrent Neural Network

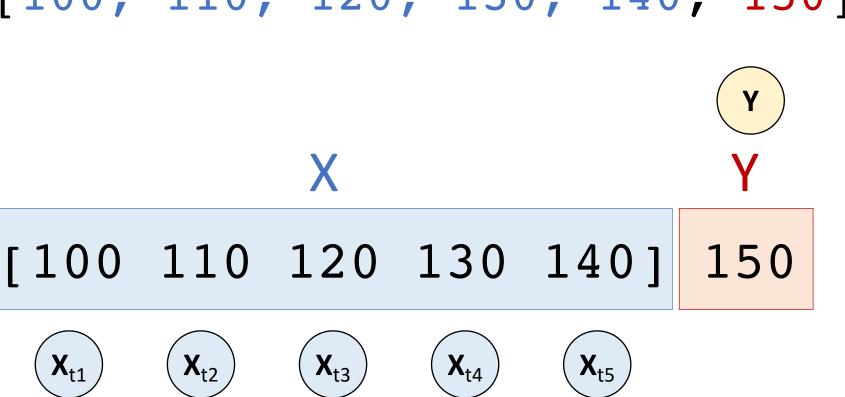


Long Short Term Memory (LSTM) for Time Series Forecasting



Time Series Data

[100, 110, 120, 130, 140, 150]

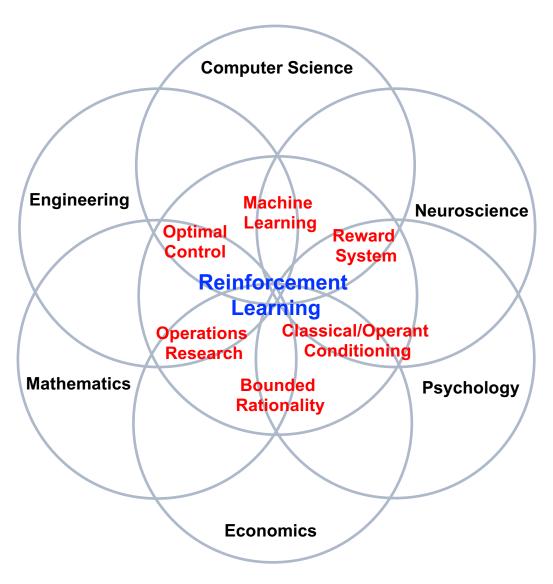


Time Series Data

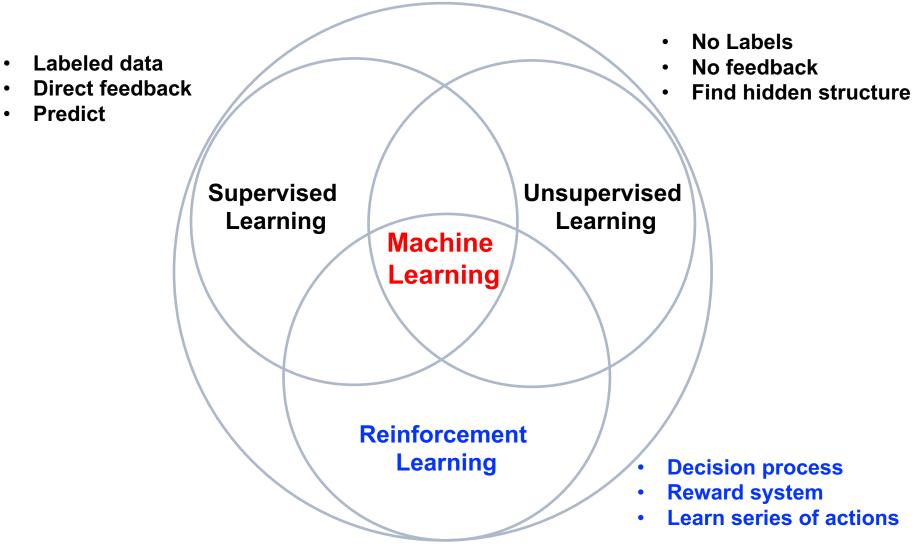
```
[10, 20, 30, 40, 50, 60, 70, 80, 90]
```

	X		Y
[10	20	30]	40
[20	30	40]	50
[30	40	50]	60
[40	50	60]	70
[50	60	70]	80
[60	70	80]	90

Reinforcement Learning (RL)



Branches of Machine Learning (ML) Reinforcement Learning (RL)

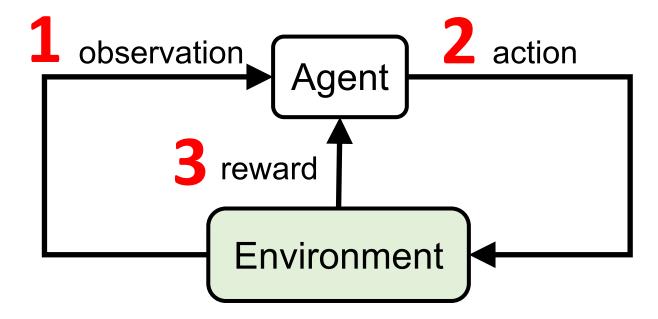


Reinforcement Learning (DL)

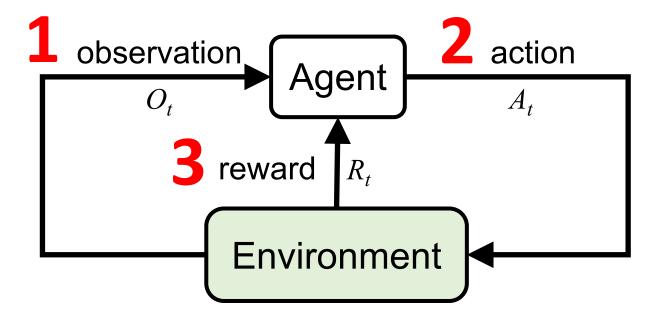
Agent

Environment

Reinforcement Learning (DL)

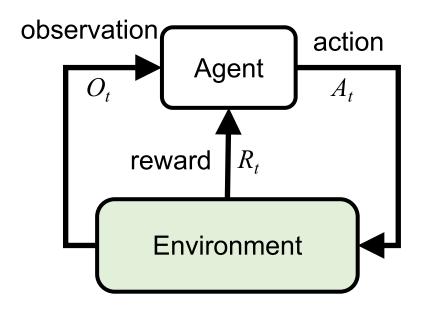


Reinforcement Learning (DL)



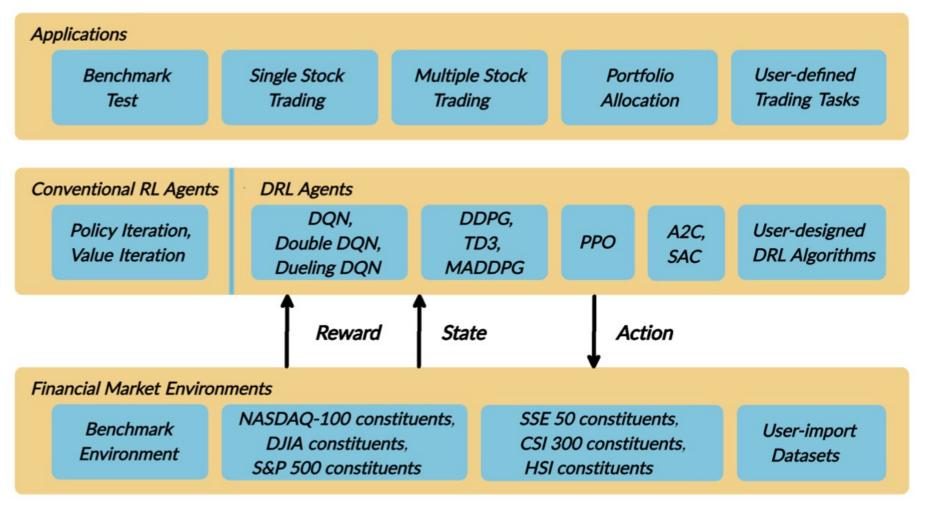
Agent and Environment

- At each step t the agent:
 - Executes action A_t
 - Receives observation O_t
 - Receives scalar reward R_t
- The environment:
 - Receives action A_t
 - Emits observation O_{t+1}
 - Emits scalar reward R_{t+1}
- t increments at env. step



FinRL:

A Deep Reinforcement Learning Library for Automated Stock Trading in Quantitative Finance



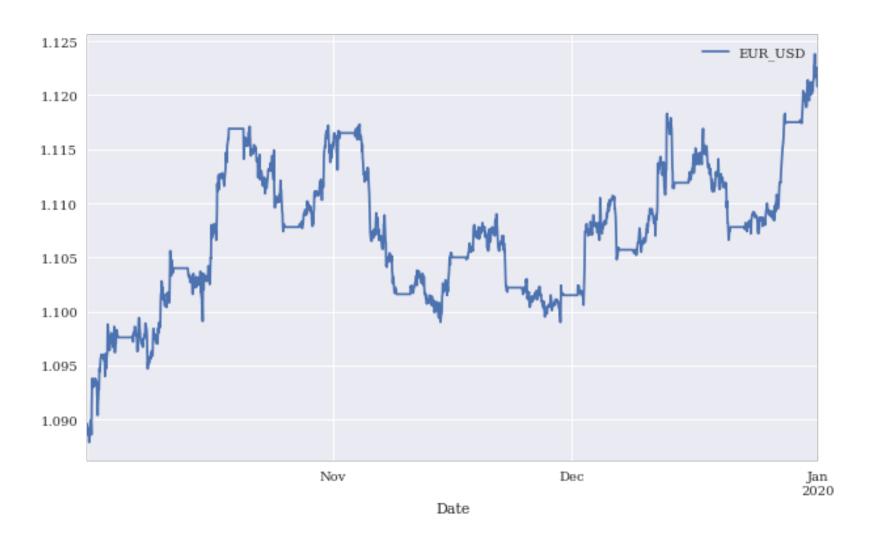


Deep Reinforcement Learning Algorithms

Algorithms	Input	Output	Туре	State-action spaces support	Finance use cases support	Features and Improvements	Advantages
DQN	States	Q-value	Value based	Discrete only	Single stock trading	Target network, experience replay	Simple and easy to use
Double DQN	States	Q-value	Value based	Discrete only	Single stock trading	Use two identical neural network models to learn	Reduce overestimations
Dueling DQN	States	Q-value	Value based	Discrete only	Single stock trading	Add a specialized dueling Q head	Better differentiate actions, improves the learning
DDPG	State action pair	Q-value	Actor-critic based	Continuous only	Multiple stock trading, portfolio allocation	Being deep Q-learning for continuous action spaces	Better at handling high-dimensiona continuous action spaces
A2C	State action pair	Q-value	Actor-critic based	Discrete and continuous	All use cases	Advantage function, parallel gradients updating	Stable, cost-effective, faster and works better with large batch sizes
PPO	State action pair	Q-value	Actor-critic based	Discrete and continuous	All use cases	Clipped surrogate objective function	Improve stability, less variance, simply to implement
SAC	State action pair	Q-value	Actor-critic based	Continuous only	Multiple stock trading, portfolio allocation	Entropy regularization, exploration-exploitation trade-off	Improve stability
TD3	State action pair	Q-value	Actor-critic based	Continuous only	Multiple stock trading, portfolio allocation	Clipped double Q-Learning, delayed policy update, target policy smoothing.	Improve DDPG performance
MADDPG	State action pair	Q-value	Actor-critic based	Continuous only	Multiple stock trading, portfolio allocation	Handle multi-agent RL problem	Improve stability and performance

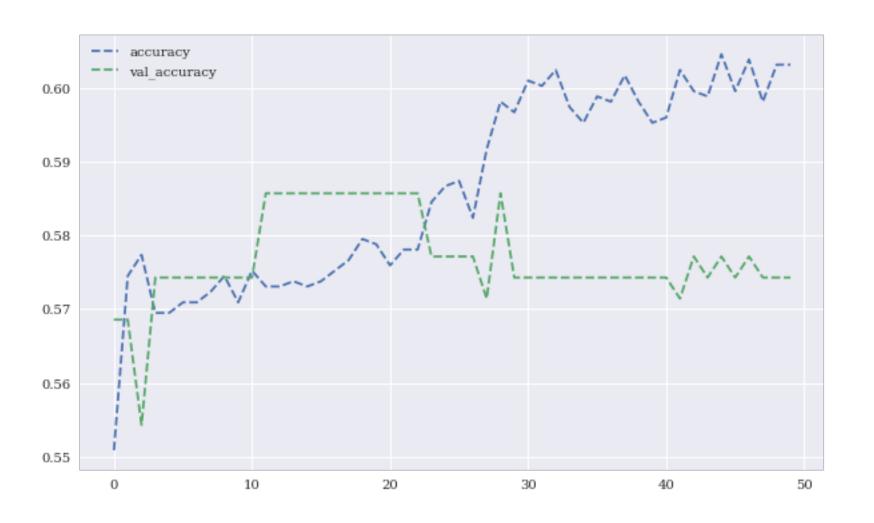
```
import os
import numpy as np
import pandas as pd
from pylab import plt, mpl
plt.style.use('seaborn')
mpl.rcParams['savefig.dpi'] = 300
mpl.rcParams['font.family'] = 'serif'
pd.set option('precision', 4)
np.set printoptions(suppress=True, precision=4)
os.environ['PYTHONHASHSEED'] = '0'
url = 'http://hilpisch.com/aiif eikon id eur usd.csv'
symbol = 'EUR USD'
raw = pd.read csv(url, index col=0, parse dates=True)
raw.head()
```

Mid-closing prices for EUR/USD (intraday)

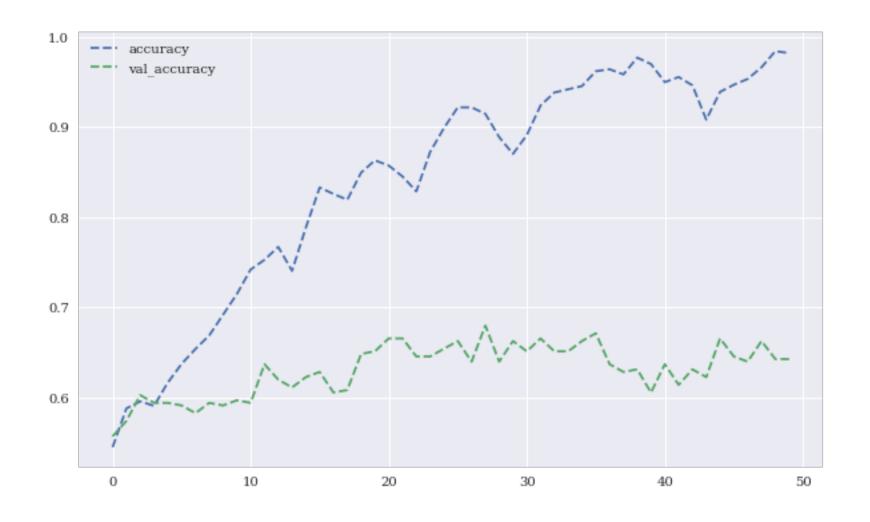


```
optimizer = Adam(lr=0.001)
def create model(hl=1, hu=128, optimizer=optimizer):
  model = Sequential()
  model.add(Dense(hu, input dim=len(cols),
               activation='relu'))
  for in range(hl):
     model.add(Dense(hu, activation='relu'))
  model.add(Dense(1, activation='sigmoid'))
  model.compile(loss='binary crossentropy',
                  optimizer=optimizer,
                 metrics=['accuracy'])
  return model
set seeds()
model = create model(hl=1, hu=128)
model.summary()
```

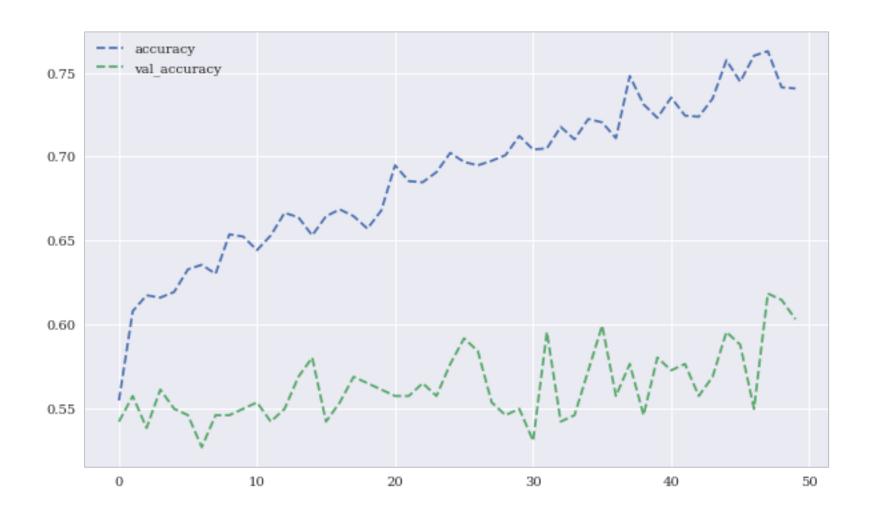
Training and validation accuracy values



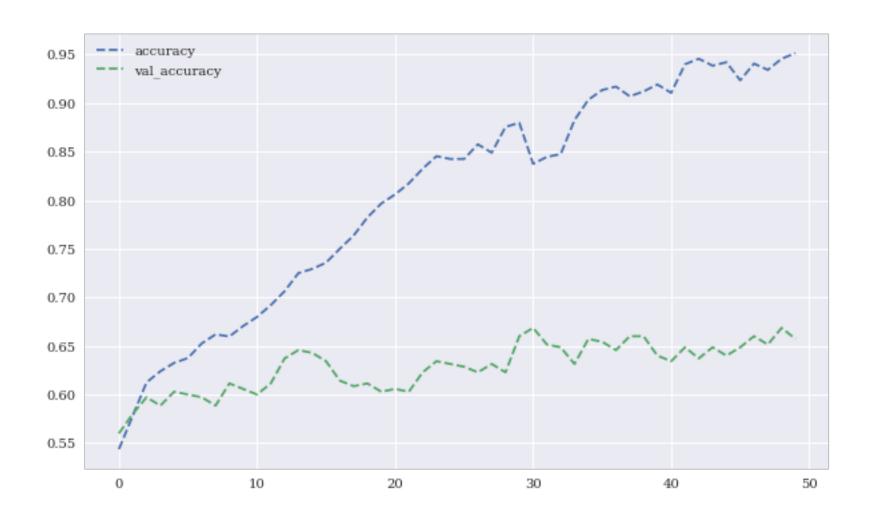
Training and validation accuracy values (normalized features data)



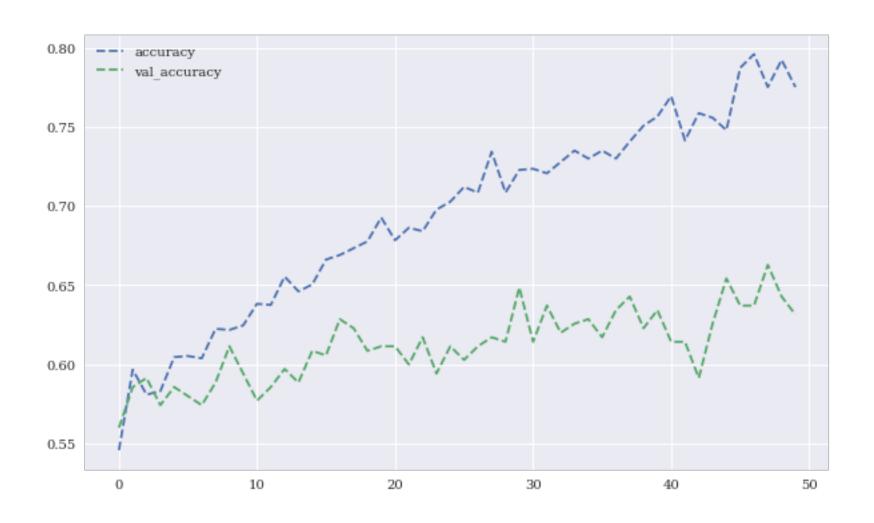
Training and validation accuracy values (with dropout)



Training and validation accuracy values (with regularization)



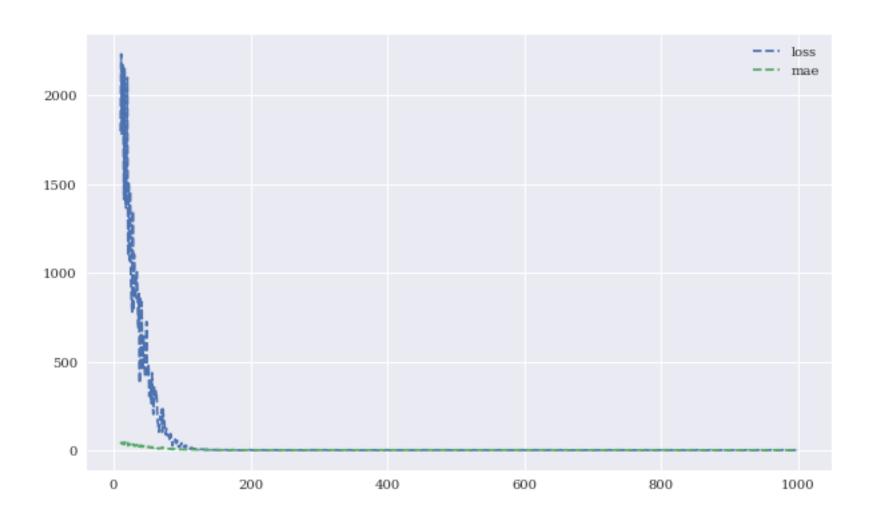
Training and validation accuracy values (with dropout and regularization)



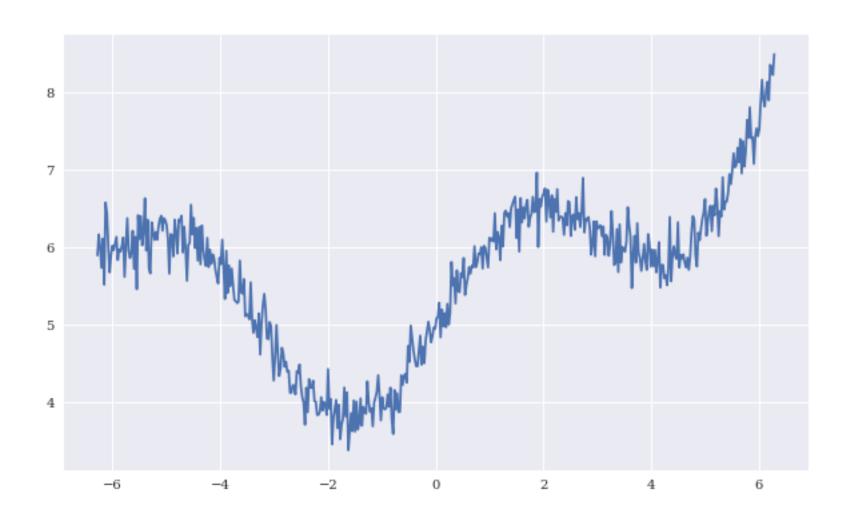
```
from keras.models import Sequential
from keras.layers import SimpleRNN, LSTM, Dense
model = Sequential()
model.add(SimpleRNN(100, activation='relu',
          input shape=(lags, 1)))
model.add(Dense(1, activation='linear'))
model.compile(optimizer='adagrad', loss='mse',
               metrics=['mae'])
model.summary()
```

model.fit(g, epochs=1000, steps per epoch=5, verbose=False)

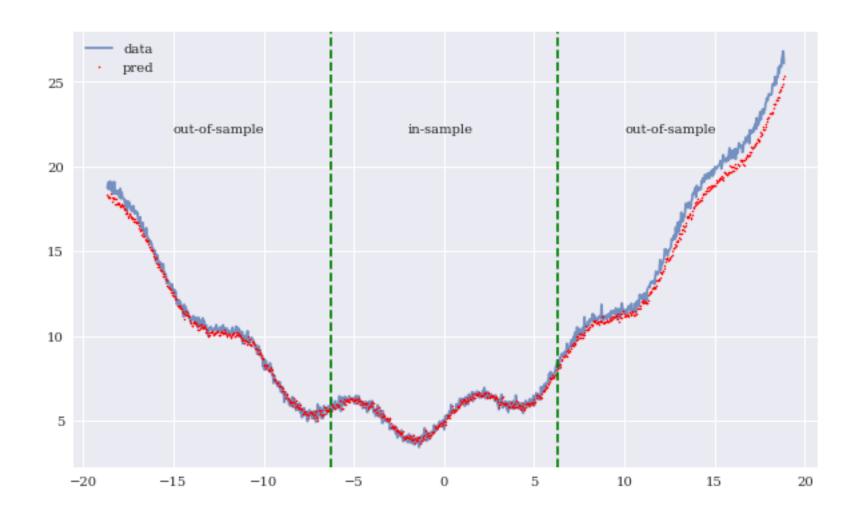
Performance metrics during RNN training



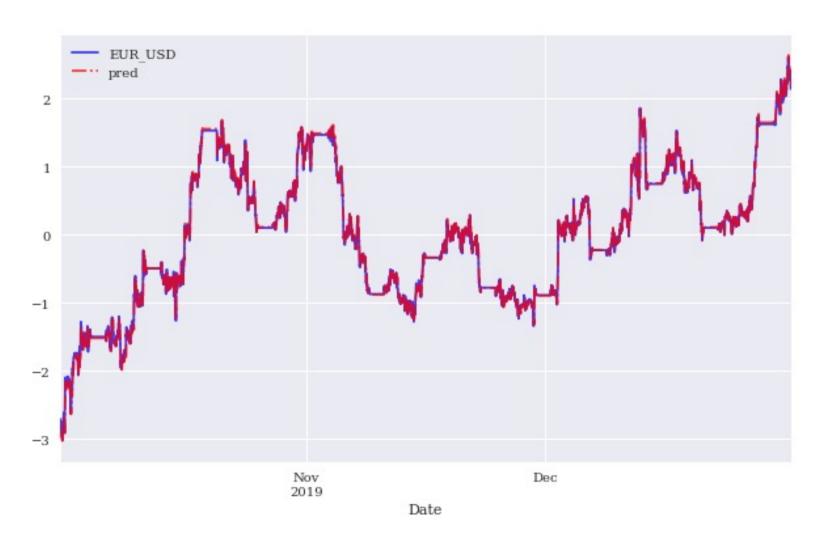
Sample sequence data



in-sample and out-of-sample predictions of the RNN



In-sample prediction for financial price series by the RNN (whole data set)

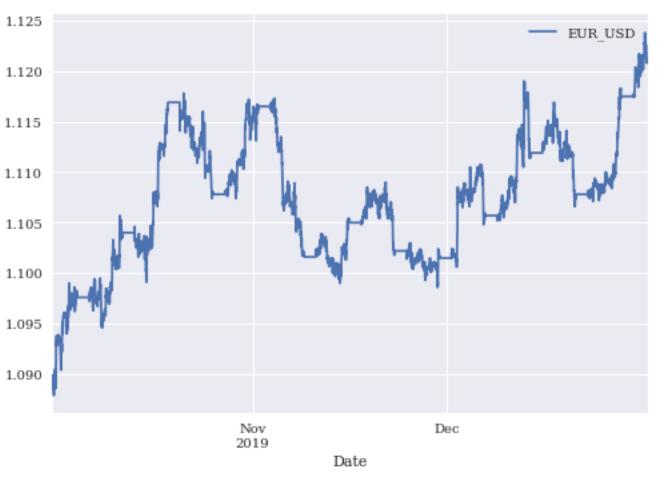


In-sample prediction for financial price series by the RNN (data sub-set)



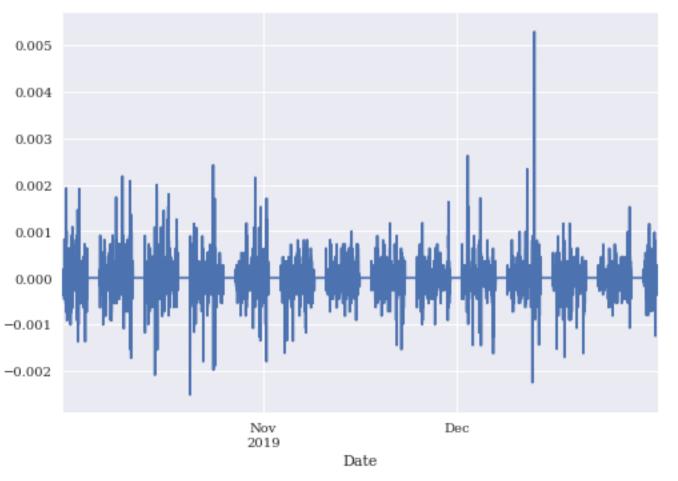
Financial Price Series

data = generate_data()
data.plot()



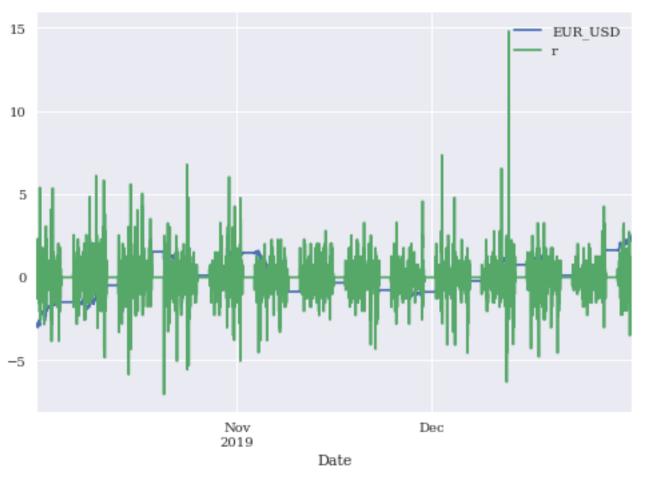
Financial Return Series

```
data['r'] = np.log(data / data.shift(1))
data['r'].plot()
```

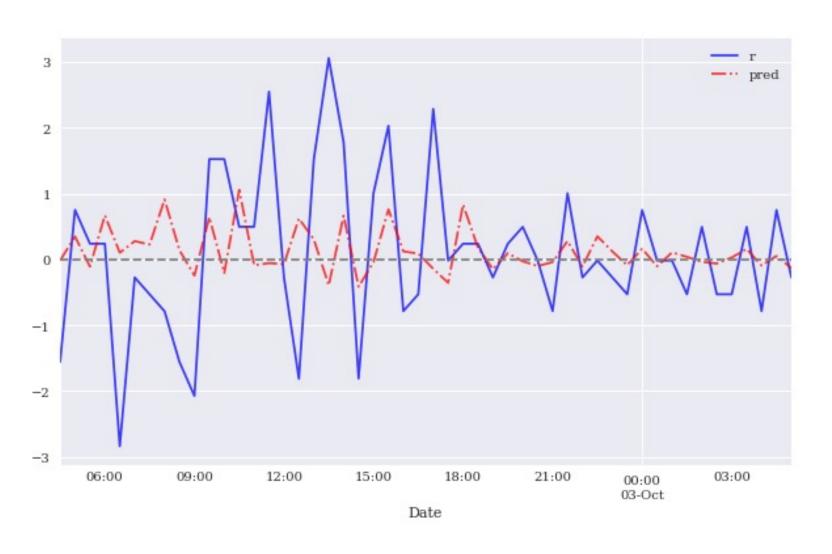


Financial Price and Return Normalization Series

```
data.dropna(inplace=True)
data = (data - data.mean()) / data.std()
data.plot()
```

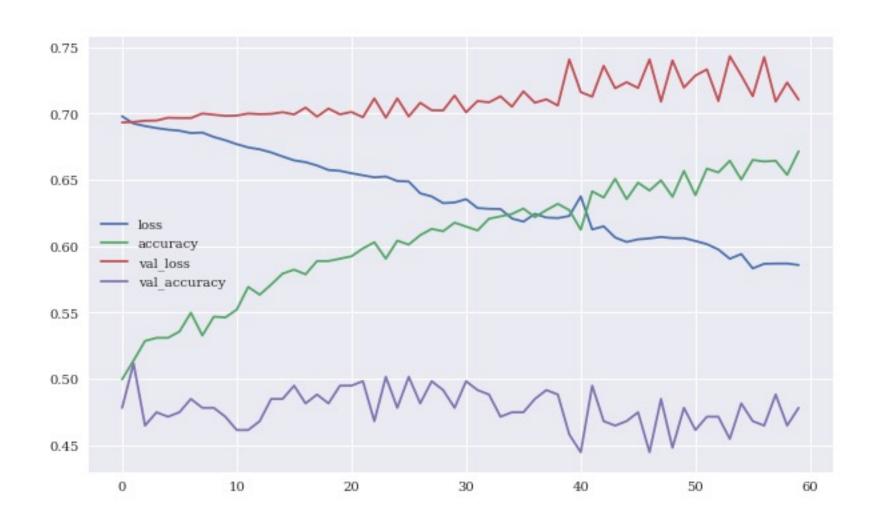


In-sample prediction for financial return series by the RNN (data sub-set)



```
model = Sequential()
model.add(Conv1D(filters=96, kernel size=5,
                  activation='relu',
                  input shape=(len(cols), 1)))
model.add(Flatten())
model.add(Dense(10, activation='relu'))
model.add(Dense(1, activation='sigmoid'))
model.compile(optimizer='adam',
               loss='binary crossentropy',
               metrics=['accuracy'])
model.summary()
model.fit(np.atleast 3d(train[cols]), train['d'],
          epochs=60, batch size=48, verbose=False,
          validation split=0.15, shuffle=False)
```

Performance metrics for the training and validation of the CNN



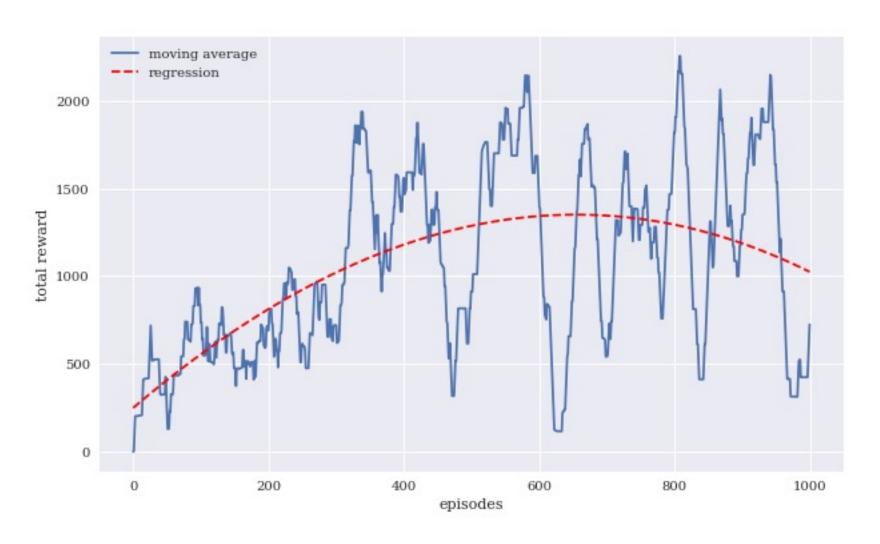
Gross performance of passive benchmark investment and CNN strategy (before/after transaction costs)



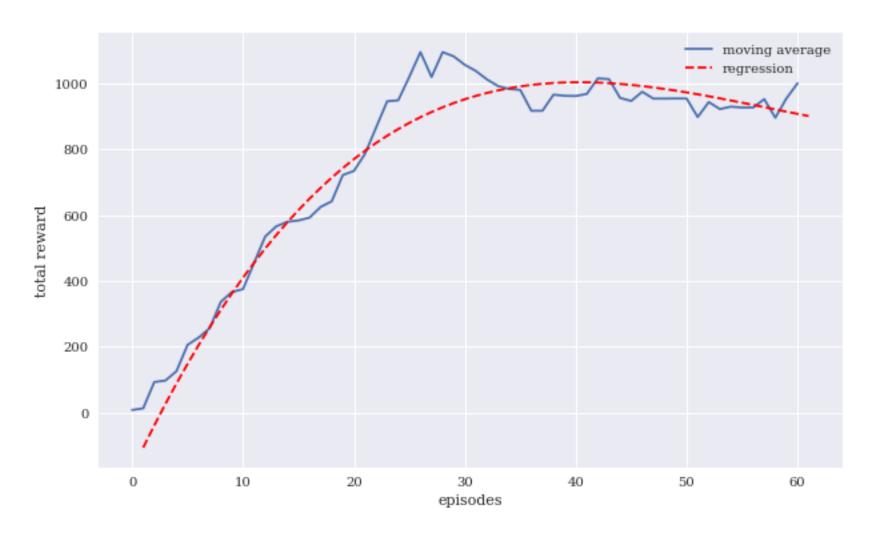
Reinforcement Learning in Finance

- Simple Learning
- DNN Learning
- Q Learning
- Finance Environment
- Improved Finance Environment
- Improved Financial QL Agent

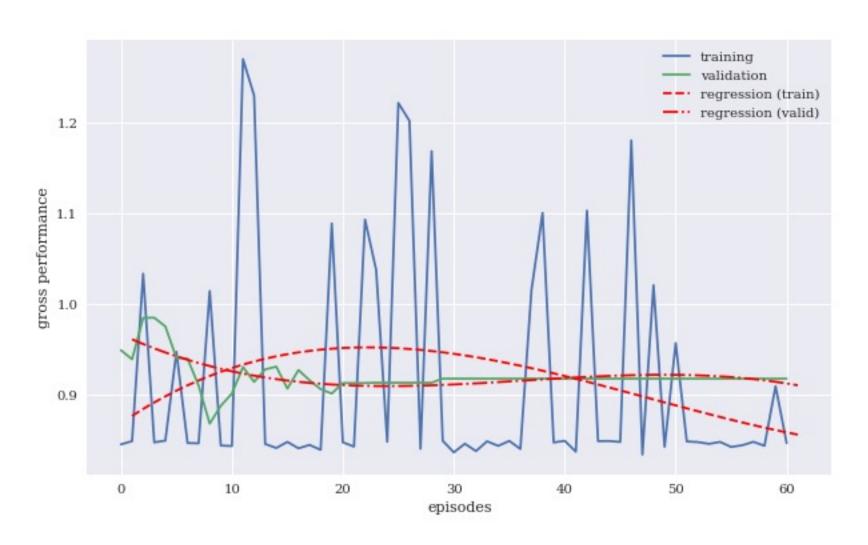
Average total rewards of DQLAgent for CartPole



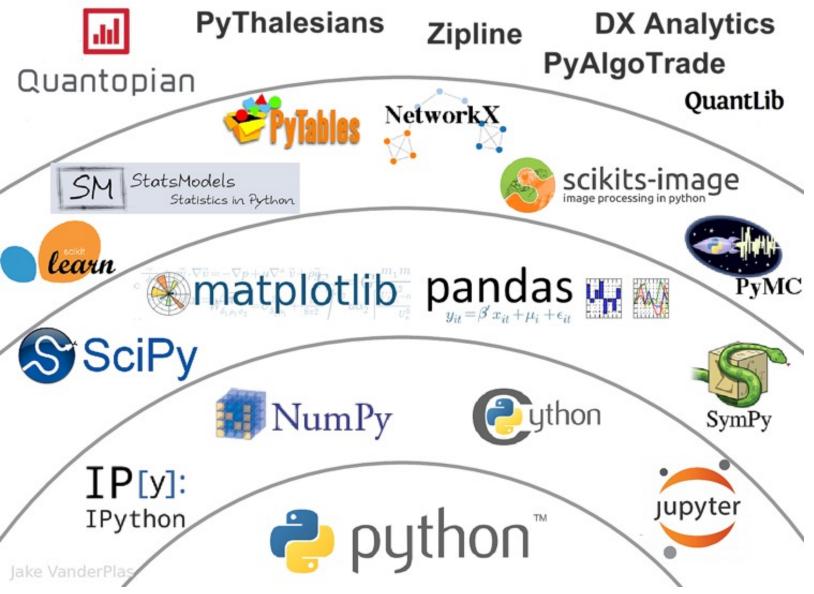
Average total rewards of DQLAgent for Finance



Training and validation performance of the FQLAgent per episode



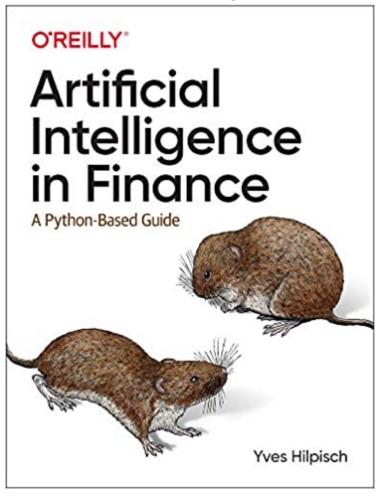
The Quant Finance PyData Stack



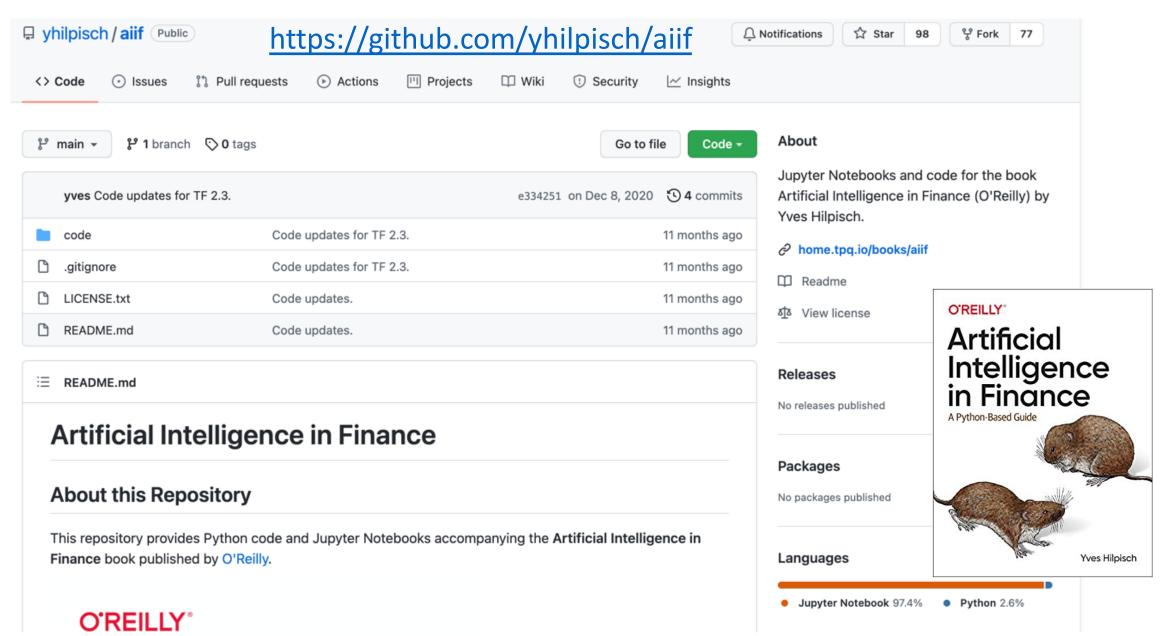
Yves Hilpisch (2020),

Artificial Intelligence in Finance: A Python-Based Guide,

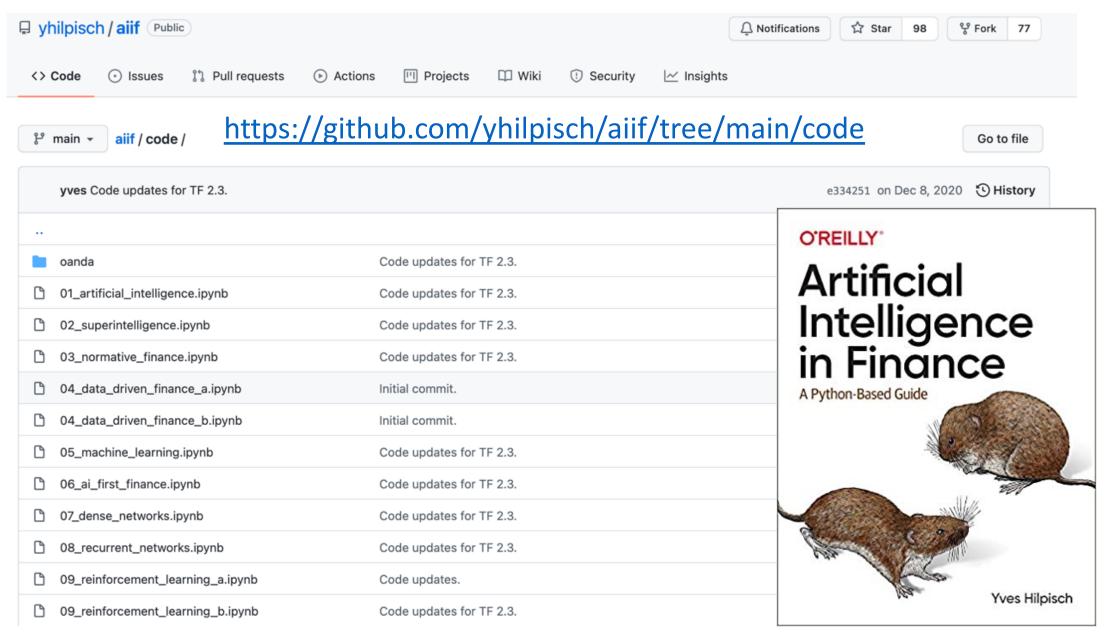
O'Reilly



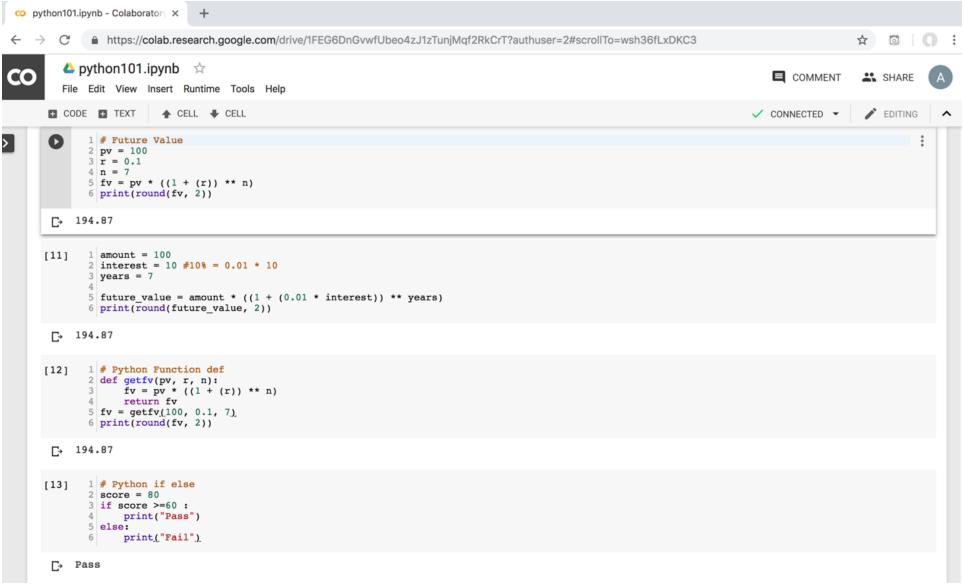
Yves Hilpisch (2020), Artificial Intelligence in Finance: A Python-Based Guide, O'Reilly



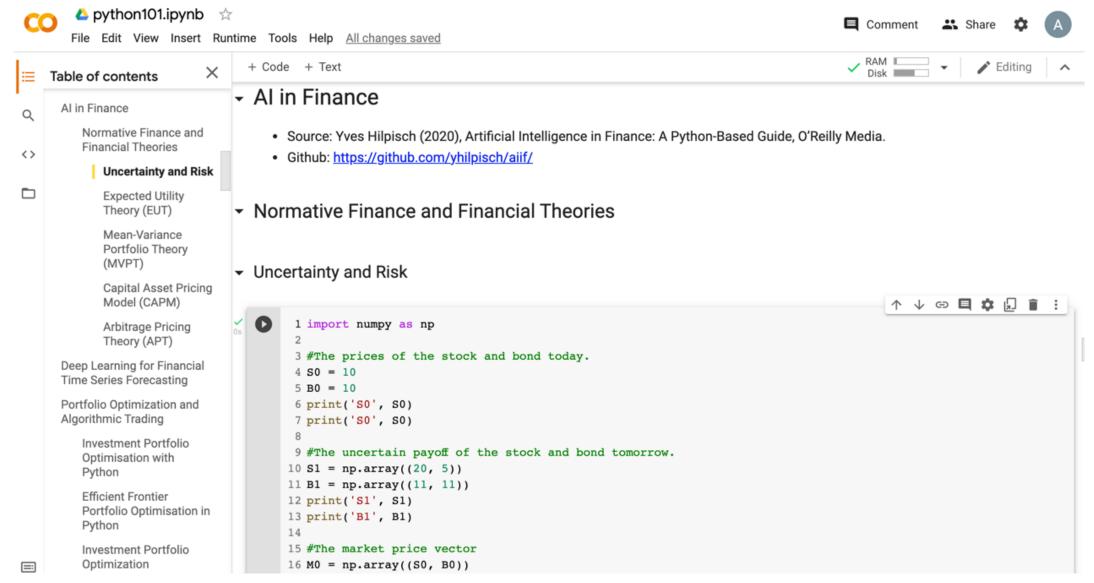
Yves Hilpisch (2020), Artificial Intelligence in Finance: A Python-Based Guide, O'Reilly

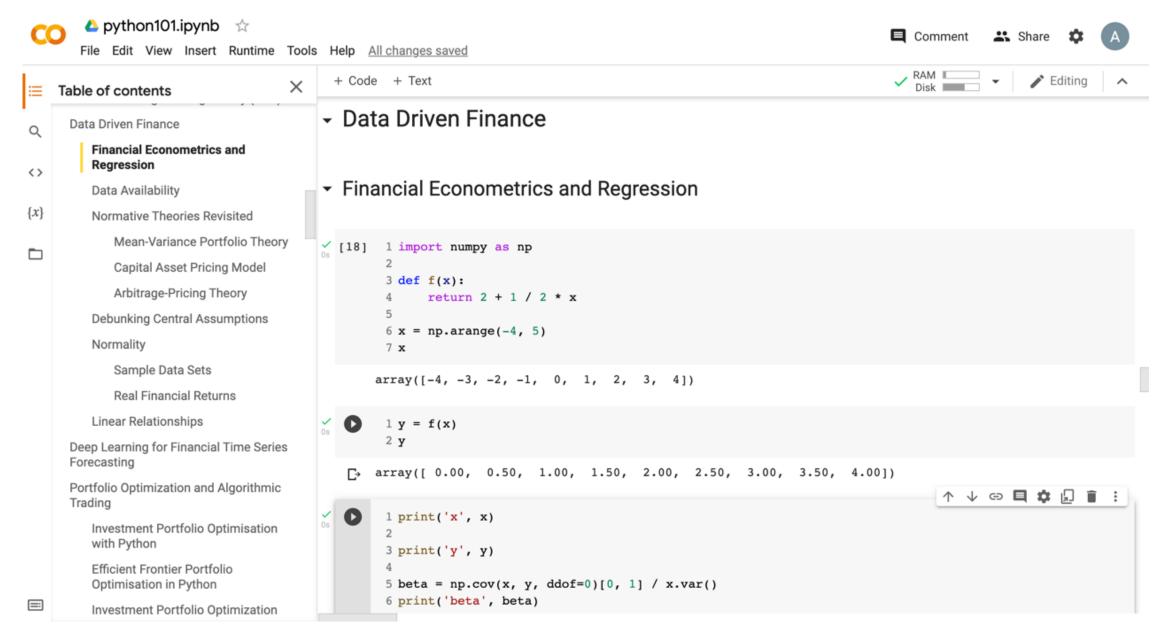


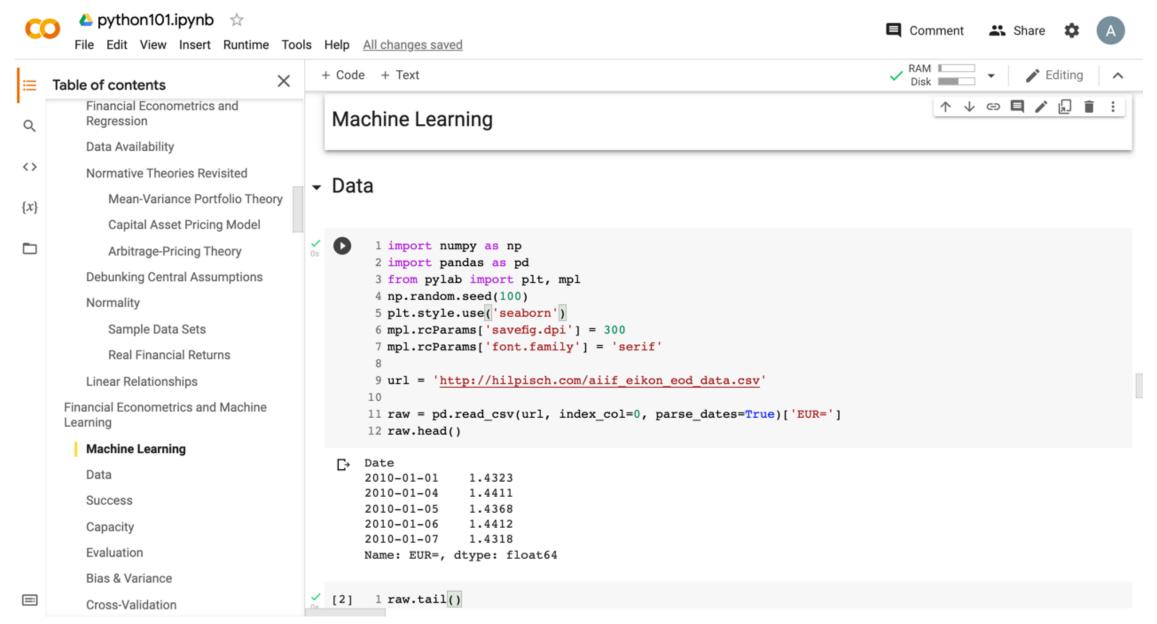
https://colab.research.google.com/drive/1FEG6DnGvwfUbeo4zJ1zTunjMqf2RkCrT

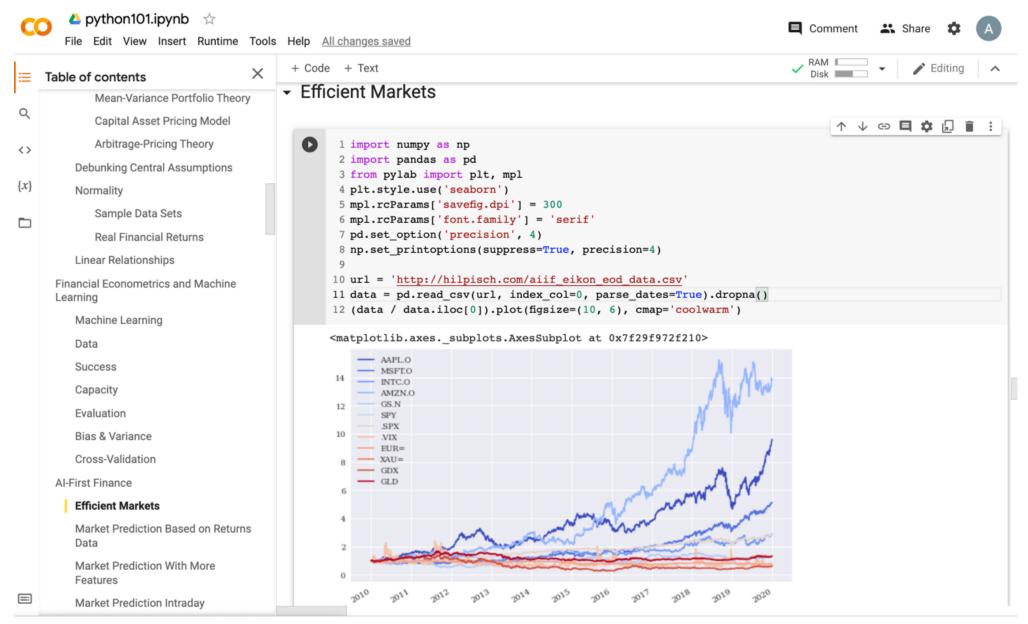


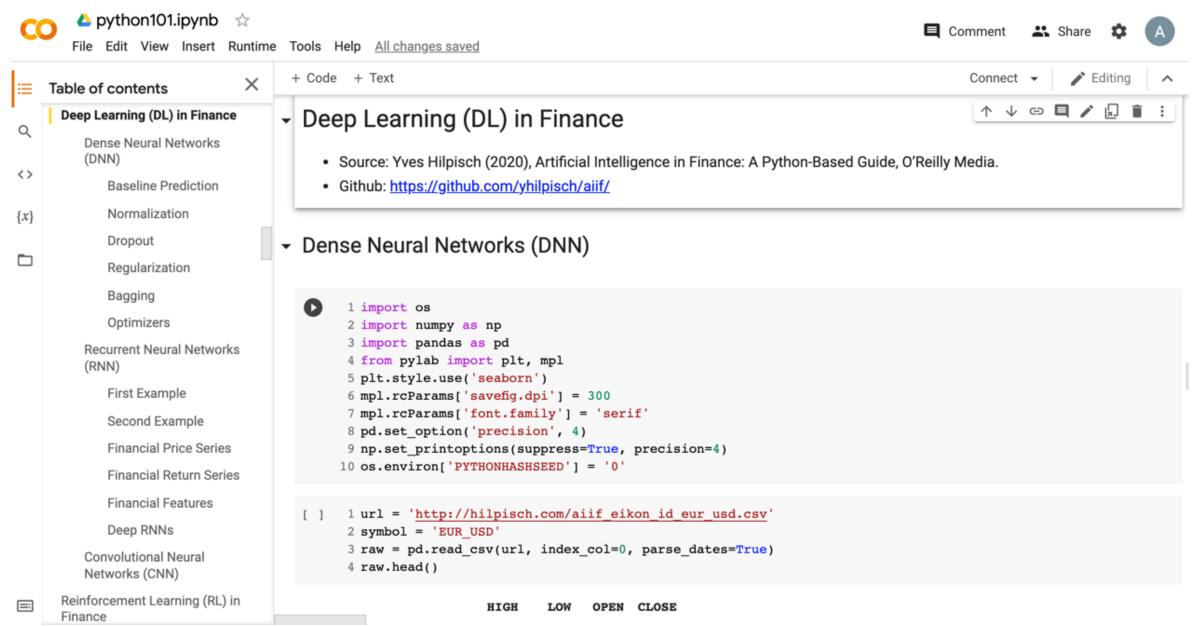
https://colab.research.google.com/drive/1FEG6DnGvwfUbeo4zJ1zTunjMqf2RkCrT

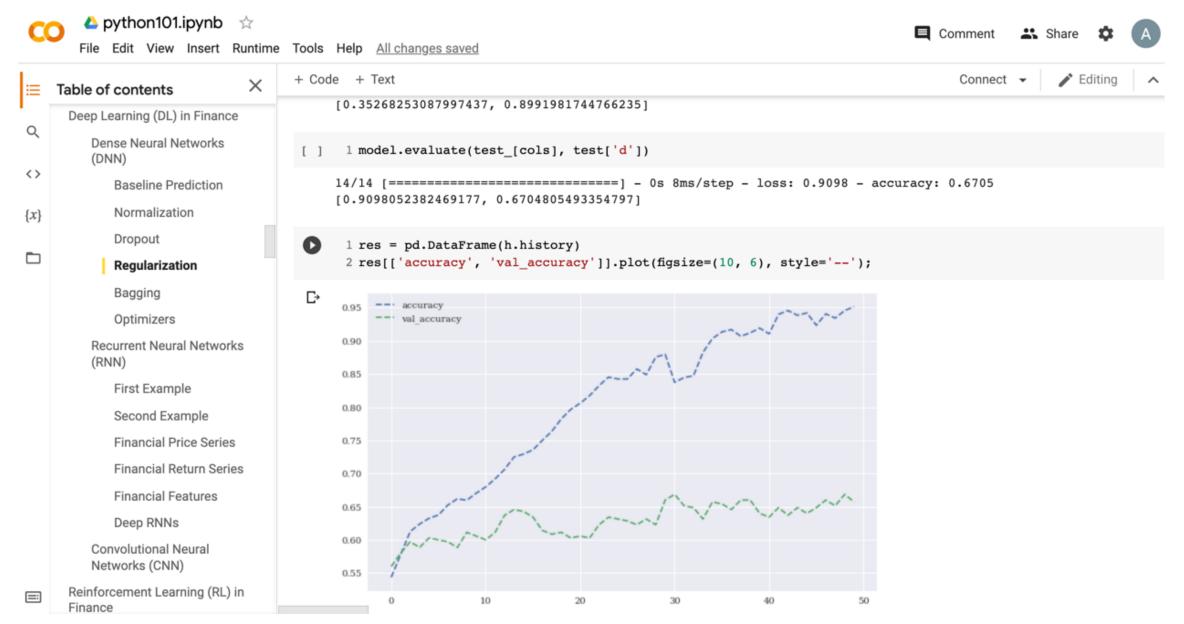


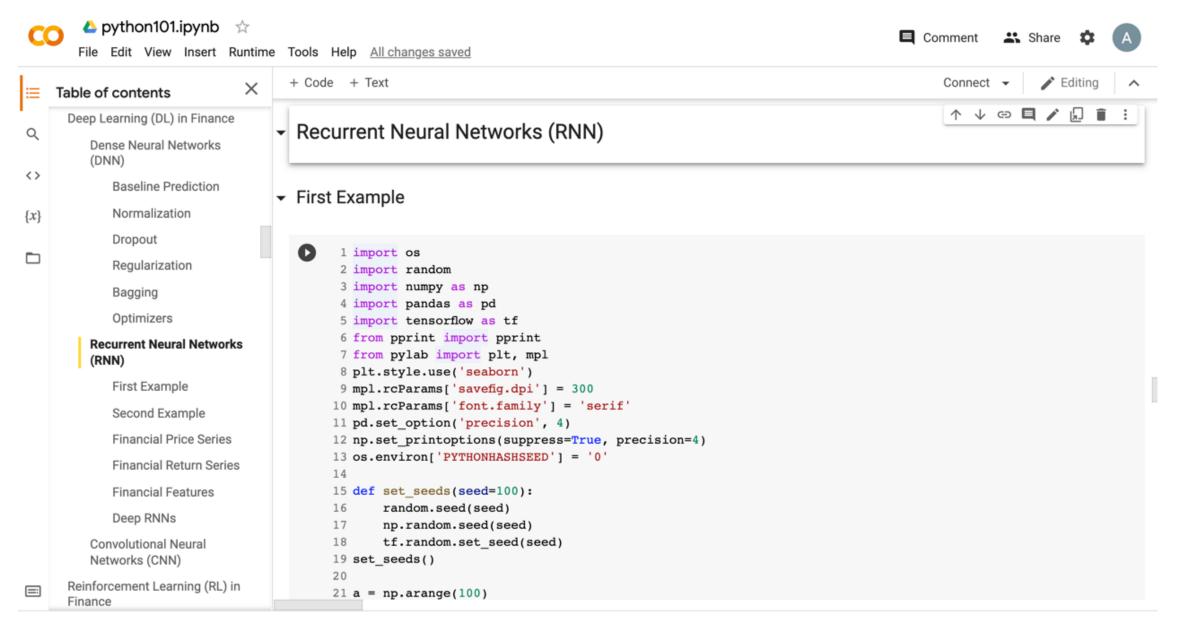


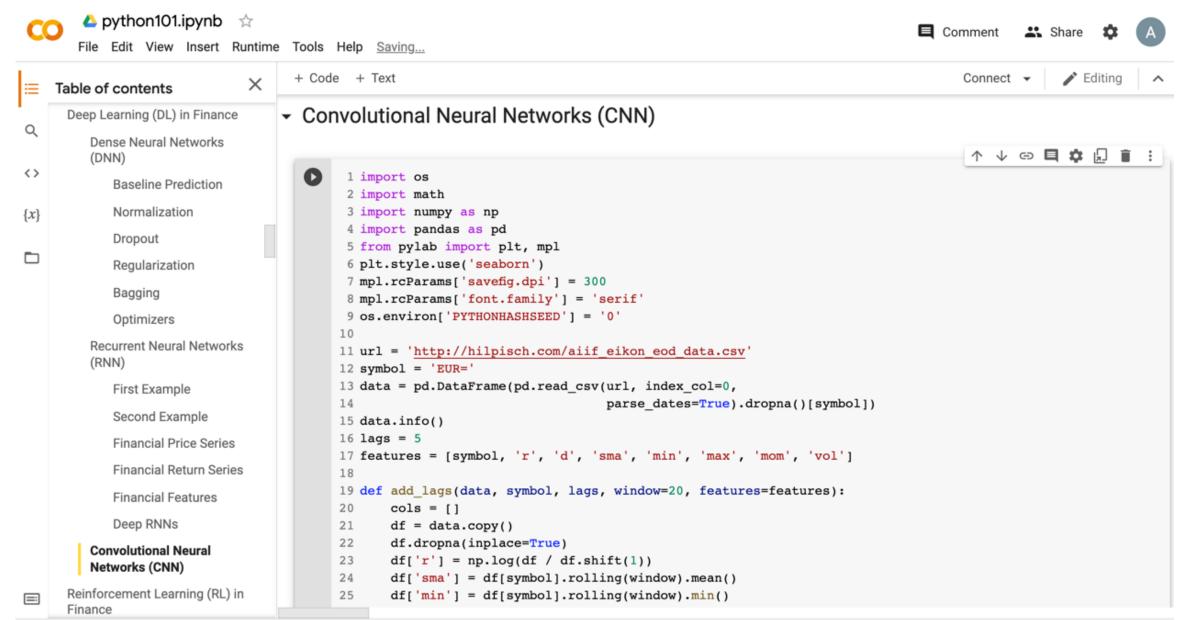


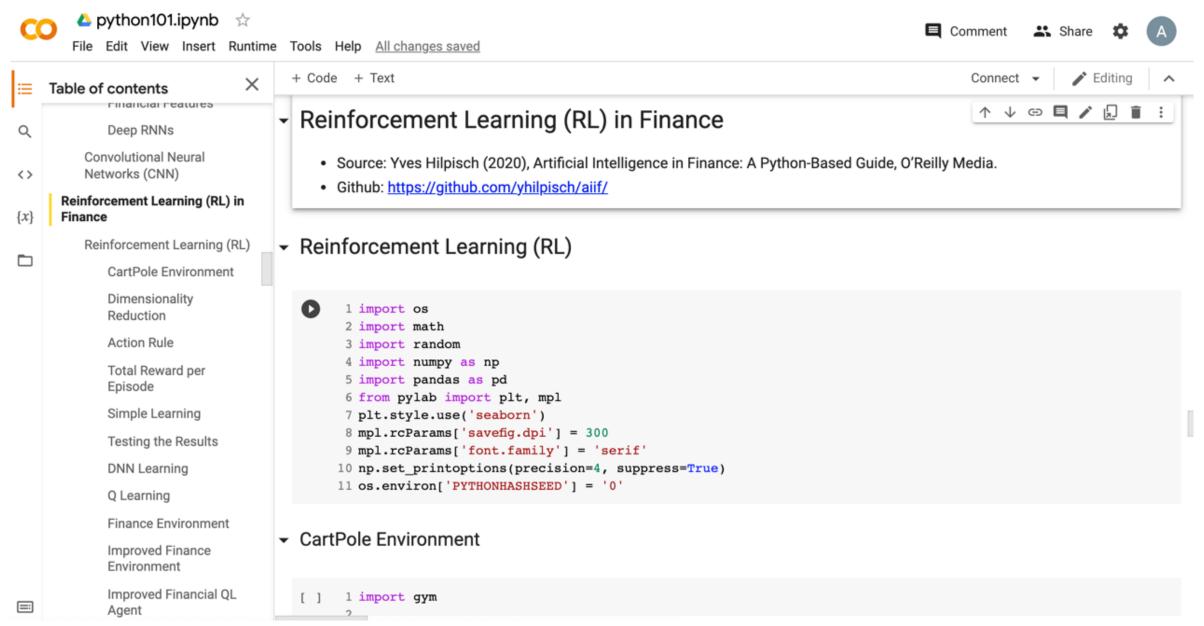


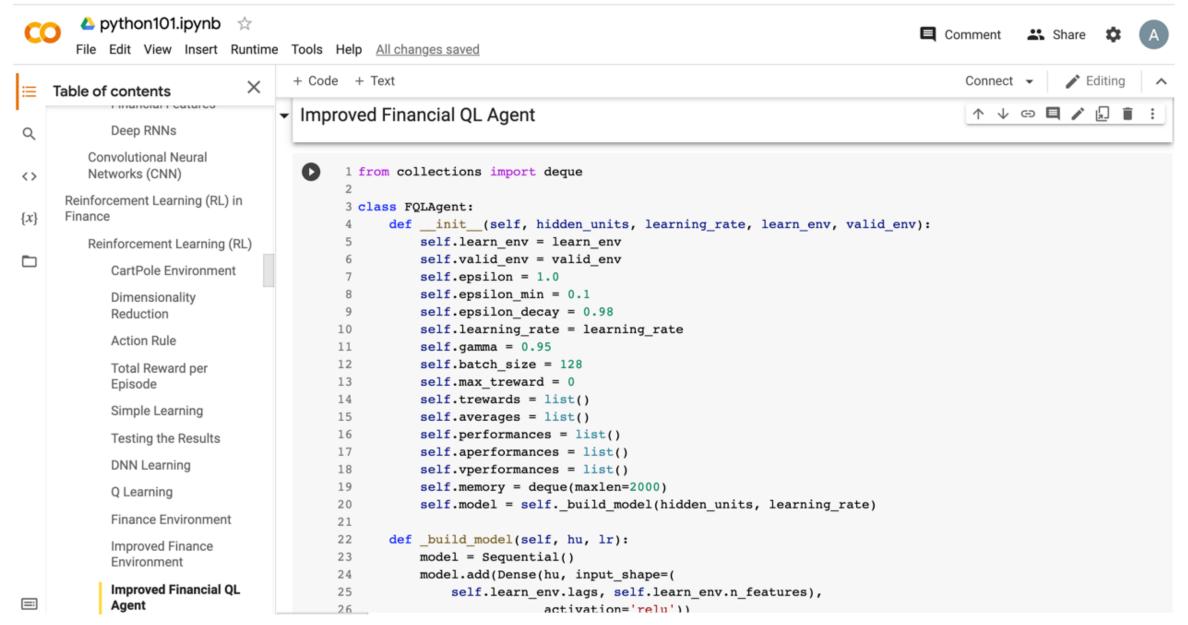












Summary

- Deep Learning (DL) in Finance
 - Dense Neural Networks (DNN)
 - Recurrent Neural Networks (RNN)
 - Convolutional Neural Networks (CNN)
- Reinforcement Learning (RL) in Finance
 - Q Learning (QL)
 - Improved Finance Environment
 - Improved Financial QL Agent

References

- Yves Hilpisch (2020), Artificial Intelligence in Finance: A Python-Based Guide, O'Reilly Media, https://github.com/yhilpisch/aiif.
- Aurélien Géron (2019), Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, 2nd Edition, O'Reilly Media.
- Ahmet Murat Ozbayoglu, Mehmet Ugur Gudelek, and Omer Berat Sezer (2020). "Deep learning for financial applications: A survey." Applied Soft Computing (2020): 106384.
- Omer Berat Sezer, Mehmet Ugur Gudelek, and Ahmet Murat Ozbayoglu (2020), "Financial time series forecasting with deep learning: A systematic literature review: 2005–2019." Applied Soft Computing 90 (2020): 106181.
- Min-Yuh Day (2022), Python 101, https://tinyurl.com/aintpupython101