Neural Network and Deep Neural Network Learning:

**1. Online Courses:**

[Deep Learning in Python (DataCamp)](https://www.datacamp.com/courses/deep-learning-in-python)

Deep Learning Specification (Couresa , Ng. Andrew) highly recommended

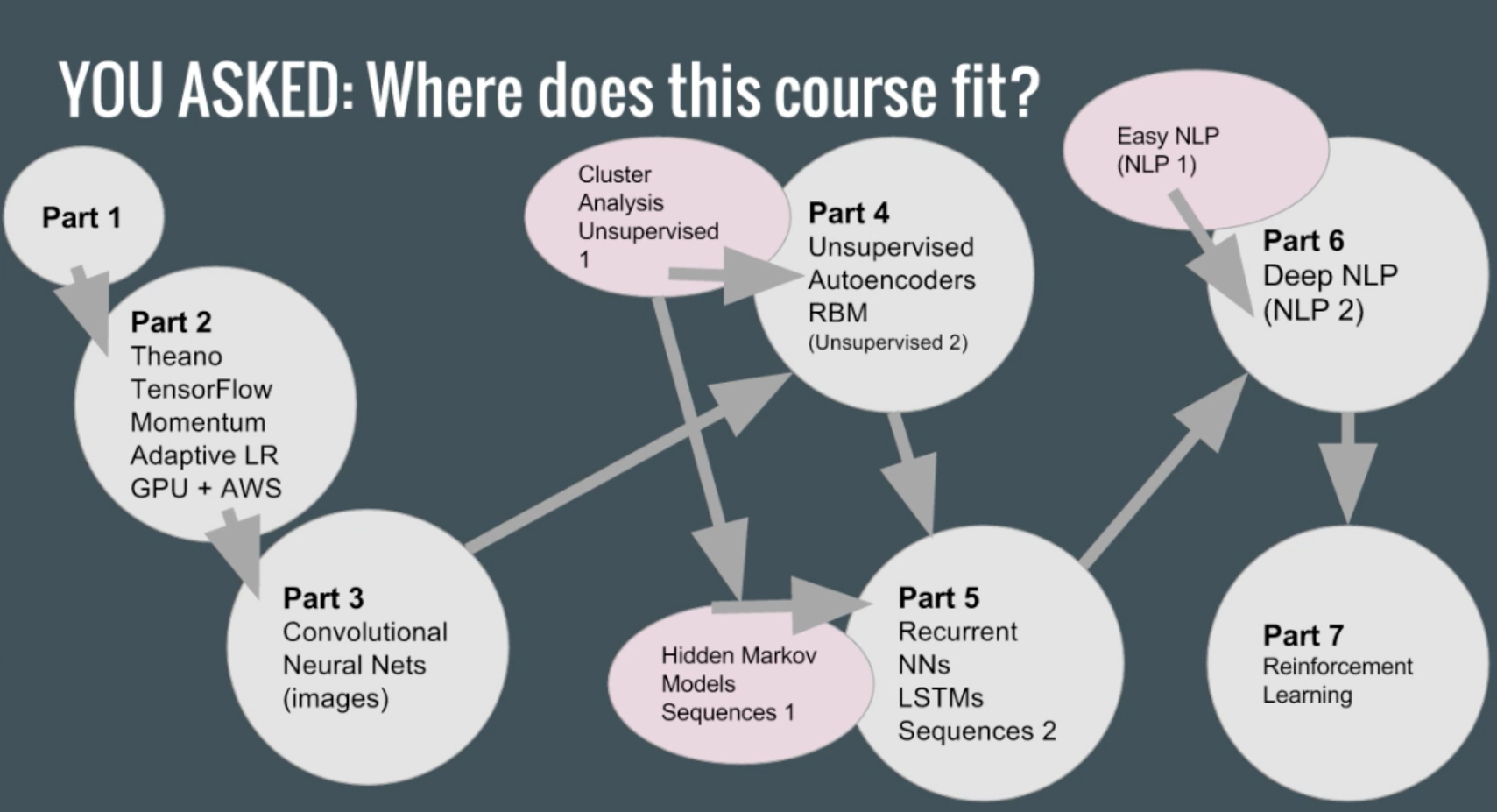
Machine Learning Nano Degree (Udacity) recommended

Deep Learning Nano Degree (Udacity) recommended

Deep Learning Foundation (Udacity , free) recommended

[Machine Learning Specialization (Coursera, University of Washington)](https://www.coursera.org/specializations/machine-learning)

**All following courses from Lazy-programming in Udemy:** highly recommended



Deep Learning Prerequisites: Linear Regression in Python, part 0

Deep Learning Prerequisites: Logistic Regression in Python, part 0

Data Science: Deep Learning in Python, part 1

Data Science: Practical Deep Learning in Theano + TensorFlow, part 2

Deep Learning: Convolutional Neural Networks in Python part 3

Unsupervised Deep Learning in Python, part 4

Deep Learning: Recurrent Neural Networks in Python, part 5

Natural Language Processing with Deep Learning in Python, part 6

Deep Learning: GANs and Variational Autoencoders

Artificial Intelligence: Reinforcement Learning in Python

Advanced AI: Deep Reinforcement Learning in Python, part 7

**Other useful machine learning courses from lazy-programming:**

Data Science: Supervised Machine Learning in Python

Cluster Analysis and Unsupervised Machine Learning in Python

Unsupervised Machine Learning Hidden Markov Models in Python

Ensemble Machine Learning in Python: Random Forest, AdaBoost

Bayesian Machine Learning in Python: A/B Testing

Data Science: Natural Language Processing (NLP) in Python

**Stanford CS231 CNN course:**

[http://cs231n.stanford.edu/](http://cs231n.stanford.edu/syllabus.html)

<https://www.youtube.com/watch?v=vT1JzLTH4G4&list=PL3FW7Lu3i5JvHM8ljYj-zLfQRF3EO8sYv>

**2. Foundation:**

**Logistic Regression:**

<http://dataaspirant.com/2017/03/02/how-logistic-regression-model-works/>

**Multi-layer NN:**

[Introduction to Multi-Layer Perceptrons (Feedforward Neural Networks)](http://www.iro.umontreal.ca/~pift6266/H10/notes/mlp.html)

**Gradient Descent:**

[Sebastian Ruder: An overview of gradient descent optimization algorithms:](http://ruder.io/optimizing-gradient-descent/index.html#downpoursgd)

[Introduction to Gradient-Based Learning](http://www.iro.umontreal.ca/~pift6266/H10/notes/gradient.html)

**Back propagation:**

<https://www.ics.uci.edu/~pjsadows/notes.pdf>

<https://deepnotes.io/softmax-crossentropy>

https://medium.com/@karpathy/yes-you-should-understand-backprop-e2f06eab496b

**Activation Function:**

<https://theclevermachine.wordpress.com/tag/tanh-function/>

[Transfer Function](http://nn.readthedocs.io/en/rtd/transfer/)

**Batch normalization:**

<https://www.quora.com/Why-does-batch-normalization-help>

**3. Online Materials:**

**Deep learning tutorials:**

<https://deepnotes.io/>

http://deeplearning.net/tutorial/

**Keras:**

Dropout regularization in deep learning models with Keras:

<https://machinelearningmastery.com/dropout-regularization-deep-learning-models-keras/>

How to checkpoint deep learning models in Keras

<https://machinelearningmastery.com/check-point-deep-learning-models-keras/>

**Project/example tutorials:**

<https://www.kaggle.com/arthurtok/spooky-nlp-and-topic-modelling-tutorial>

<http://www.jianshu.com/p/482b32716bbe>

**Papers**

The Marginal Value of Adaptive Gradient Methods in Machine Learning

<https://arxiv.org/abs/1705.08292>

Asynchronous Stochastic Gradient Descent with Delay Compensation for Distributed Deep Learning

<https://arxiv.org/abs/1609.08326>

Asynchronous Stochastic Gradient Descent with Variance Reduction for Non-Convex Optimization

<https://arxiv.org/abs/1604.03584>

Large Scale Distributed Deep Networks

<https://static.googleusercontent.com/media/research.google.com/en//archive/large_deep_networks_nips2012.pdf>

Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift

Sergey Ioffe, Christian Szegedy

<https://arxiv.org/abs/1502.03167>

Xavier (Glorot) Normal Initializer

<http://jmlr.org/proceedings/papers/v9/glorot10a/glorot10a.pdf>

Delving Deep into Rectifiers: Surpassing Human-Level Performance on ImageNet Classification

<http://arxiv.org/abs/1502.01852>

**Other articles:**

<https://www.kdnuggets.com/2017/11/ng-deep-learning-specialization-21-lessons.html>

<https://www.kdnuggets.com/2017/11/unit-test-machine-learning-code.html>

<https://www.kdnuggets.com/2017/11/forget-for-loop-data-science-code-vectorization.html>

<https://www.kdnuggets.com/2017/03/software-engineering-vs-machine-learning-concepts.html>

<https://www.analyticsvidhya.com/blog/2016/01/complete-tutorial-ridge-lasso-regression-python/>

<https://www.kdnuggets.com/2017/11/understanding-deep-convolutional-neural-networks-tensorflow-keras.html>

<https://www.kdnuggets.com/2017/10/understanding-machine-learning-algorithms.html>

<https://www.kdnuggets.com/2017/10/top-10-machine-learning-algorithms-beginners.html>

<https://www.kdnuggets.com/2017/08/machine-learning-algorithms-concise-technical-overview-part-1.html>

<https://www.kdnuggets.com/2017/11/3-different-types-machine-learning.html>

<https://www.kdnuggets.com/2017/10/alphago-zero-biggest-ai-advance.html>

<https://www.kdnuggets.com/2017/12/interview-rich-sutton-reinforcement-learning.html>