

## Books

1. Artificial Intelligence: A Modern Approach. *Stuart Russell* and *Peter Norvig* [AIMA]
2. Deep Learning. *Ian Goodfellow*, *Yoshua Bengio* and *Aaron Courville* [DL]

## Section B [External]

[AIMA]

### 18.3 LEARNING DECISION TREES

18.3.1 The decision tree representation

18.3.2 Expressiveness of decision trees

18.3.3 Inducing decision trees from examples

18.3.4 Choosing attribute tests

Confusion matrix, precision, recall, f1-score

[https://en.wikipedia.org/wiki/Confusion\\_matrix](https://en.wikipedia.org/wiki/Confusion_matrix)

### 18.4 EVALUATING AND CHOOSING THE BEST HYPOTHESIS

18.4.1 Model selection: Complexity versus goodness of fit

18.4.2 From error rates to loss

18.4.3 Regularization

### 18.6 REGRESSION AND CLASSIFICATION WITH LINEAR MODELS

18.6.1 Univariate linear regression

Univariate polynomial regression

18.6.2 Multivariate linear regression

Lasso, Ridge, Elastic net, Sparsity

### 18.10 ENSEMBLE LEARNING

AdaBoost

Ensemble learning

[http://www.cs.man.ac.uk/~gbrown/publications/brown10ensemblelearning\\_proof.pdf](http://www.cs.man.ac.uk/~gbrown/publications/brown10ensemblelearning_proof.pdf)

[\*Theoretical Perspectives: Ensemble Diversity ... this section is not included in syllabus]

### 20.1 STATISTICAL LEARNING

#### 20.2 LEARNING WITH COMPLETE DATA

20.2.1 Maximum-likelihood parameter learning: Discrete models

20.2.3 Maximum-likelihood parameter learning: Continuous models

#### 20.3 LEARNING WITH HIDDEN VARIABLES: THE EM ALGORITHM

20.3.1 Unsupervised clustering: Learning mixtures of Gaussians

## Section A [Internal]

Dimensionality reduction, Principal component analysis [Up to Slide 13]

[https://cse.iitk.ac.in/users/piyush/courses/ml\\_autumn16/771A\\_lec11\\_slides.pdf](https://cse.iitk.ac.in/users/piyush/courses/ml_autumn16/771A_lec11_slides.pdf)

Matrix factorization and ALS [Up to Slide 14]

[https://cse.iitk.ac.in/users/piyush/courses/ml\\_autumn16/771A\\_lec14\\_slides.pdf](https://cse.iitk.ac.in/users/piyush/courses/ml_autumn16/771A_lec14_slides.pdf)

[DL]

5.1 Learning Algorithms

5.1.1 The Task, T

5.1.2 The Performance Measure, P

5.1.3 The Experience, E

6.5.1 Computational graph

6.5.2 Chain Rule of Calculus

6.5.3 Recursively Applying the Chain Rule to Obtain Backprop

Algorithm 6.1 and 6.2

Backpropagation Intuition

<http://cs231n.github.io/optimization-2/>

CNN

<http://cs231n.github.io/convolutional-networks/>

RNN, LSTM

<http://colah.github.io/posts/2015-08-Understanding-LSTMs/>

Backpropagation in Time BPTT

<http://www.wildml.com/2015/10/recurrent-neural-networks-tutorial-part-3-backpropagation-through-time-and-vanishing-gradients/>

Word2Vec

<http://mccormickml.com/2016/04/19/word2vec-tutorial-the-skip-gram-model/>

Autoencoder, Regularized, Sparse, Denoising

<https://www.jeremyjordan.me/autoencoders/>

<https://medium.com/analytics-vidhya/journey-from-principle-component-analysis-to-autoencoders-e60d066f191a>

Stacked Autoencoder and Layer-wise Pretraining

[http://ufldl.stanford.edu/wiki/index.php/Stacked\\_Autoencoders](http://ufldl.stanford.edu/wiki/index.php/Stacked_Autoencoders)

Variational Autoencoder

Optimization

8.3.1 Stochastic Gradient Descent [Algorithm 8.1]

8.3.2 Momentum [Algorithm 8.2]

8.3.3 Nesterov Momentum [Algorithm 8.3]

8.5.1 AdaGrad [Algorithm 8.4]

8.5.2 RMSProp [Algorithm 8.5, 8.6]

8.5.3 Adam [Algorithm 8.7]