

# Contents

---

<b>Contents</b> .....	3
<b>Introduction</b> .....	7
Audience for this Book.....	8
Review of Multiple-Layer Feedforward Networks.....	9
What Are Deep Belief Nets, and Why Do We Like Them?.....	12
<b>Supervised Feedforward Networks</b> .....	15
Back Propagation of Errors.....	16
SoftMax Outputs for Classification.....	20
Code for Gradient Calculation.....	26
Weight Penalties.....	36
Multi-Threading Gradient Computation.....	38
Gradient Computation with CUDA.....	45
Basic Architecture.....	45
A Simple Example.....	48
Initialization.....	53
Hidden Neuron Activation.....	58
Output Neuron Activation.....	63
SoftMax Output.....	64
Output Delta.....	66
Output Gradient.....	67
Gradient of the First Hidden Layer.....	69
Gradient of Subsequent Hidden Layers.....	72
Fetching the Gradient.....	75
Mean Squared Error by Reduction.....	78
Log Likelihood by Reduction.....	83
Putting It All Together.....	84

---

Basic Training Algorithms.....	91
Simulated Annealing for Starting Weights.....	91
Singular Value Decomposition for Optimal Output Weights	94
Stochastic Gradient Descent. ....	96
Conjugate Gradient Optimization. ....	99
<b>Restricted Boltzmann Machines.....</b>	<b>107</b>
What is a Restricted Boltzmann Machine?.....	108
Reconstruction Error. ....	110
Maximum Likelihood Training - Sort of.....	112
Contrastive Divergence. ....	115
Weight Penalties. ....	119
Encouraging Sparsity.....	120
Finding Initial Weights.....	122
Hidden Neuron Bias.....	124
Visible Neuron Bias. ....	125
Code for Reconstruction Error. ....	126
Multi-Threading Initial Weight Selection. ....	128
Stochastic Gradient Descent Basic Principles. ....	135
The Core Algorithm. ....	135
Dividing Epochs into Batches.....	136
Shuffling Epochs.....	137
Updating the Learning Rate and Momentum. ....	138
Determining Convergence.....	140
Code for Multi-Threaded RBM Training.....	142
CUDA Code for RBM Training.....	162
Initialization and Cache Line Matching.....	162
Fetching Training Cases.....	164
Visible-to-Hidden Layer.....	166
Hidden-to-Visible Layer.....	167
Gradient Length and Dot Product by Reduction.....	168
Updating the Input Bias.....	169
Updating the Hidden Neuron Bias. ....	170
Updating the Weights. ....	172

---

Putting It All Together.....	175
Timing.....	183
Updating Weights Analysis.....	184
Visible-to-Hidden Analysis.....	189
Hidden-to-Visible Analysis.....	192
Advanced Training and Future Versions.....	195
<b>Greedy Training.....</b>	<b>197</b>
Generative Sampling.....	200
<b>DEEP Operating Manual.....</b>	<b>207</b>
Menu Options.....	208
File Menu Options.....	208
Test Menu Options.....	209
Display Menu Options.....	210
Read a Database.....	211
Read MNIST Image.....	212
Read MNIST Labels.....	212
Write Activation File.....	213
Clear All Data.....	213
Model Architecture.....	214
Database Inputs and Targets.....	215
RBM Training Params.....	216
Supervised Training Params.....	220
Train.....	223
Test.....	226
Analyze.....	227
Receptive Field.....	228
Generative Sample.....	229
The DEEP.LOG File.....	232
<b>Index.....</b>	<b>239</b>