

Table of Contents

About the Author vii

About the Technical Reviewer ix

Chapter 1: Introduction..... 1

 Review of Multiple-Layer Feedforward Networks..... 2

 What Are Deep Belief Nets, and Why Do We Like Them? 5

Chapter 2: Supervised Feedforward Networks 9

 Backpropagation of Errors 9

 SoftMax Outputs for Classification 13

 Code for Gradient Calculation 18

 Weight Penalties 27

 Multithreading Gradient Computation..... 29

 Gradient Computation with CUDA 36

 Basic Architecture 37

 A Simple Example..... 39

 Initialization 43

 Hidden Neuron Activation 47

 Output Neuron Activation..... 52

 SoftMax Output..... 53

 Output Delta..... 54

 Output Gradient 55

 Gradient of the First Hidden Layer..... 57

 Gradient of Subsequent Hidden Layers 60

 Fetching the Gradient 63

 Mean Squared Error by Reduction..... 65

TABLE OF CONTENTS

Log Likelihood by Reduction.....	70
Putting It All Together	71
Basic Training Algorithms.....	77
Simulated Annealing for Starting Weights	78
Singular Value Decomposition for Optimal Output Weights	80
Stochastic Gradient Descent	82
Conjugate Gradient Optimization	84
Chapter 3: Restricted Boltzmann Machines	91
What Is a Restricted Boltzmann Machine?	91
Reconstruction Error.....	93
Maximum Likelihood Training, Sort Of	94
Contrastive Divergence	97
Weight Penalties.....	100
Encouraging Sparsity	101
Finding Initial Weights.....	103
Hidden Neuron Bias.....	105
Visible Neuron Bias.....	106
Code for Reconstruction Error	107
Multithreading Initial Weight Selection.....	109
Stochastic Gradient Descent Basic Principles	116
The Core Algorithm	116
Dividing Epochs into Batches	117
Shuffling Epochs.....	118
Updating the Learning Rate and Momentum.....	118
Determining Convergence	120
Code for Multithreaded RBM Training	121
CUDA Code for RBM Training.....	141
Initialization and Cache Line Matching.....	142
Fetching Training Cases.....	143

Visible-to-Hidden Layer	144
Hidden-to-Visible Layer	145
Gradient Length and Dot Product by Reduction.....	146
Updating the Input Bias	148
Updating the Hidden Neuron Bias	149
Updating the Weights	151
Putting It All Together	153
Timing.....	160
Updating Weights Analysis	161
Visible-to-Hidden Analysis.....	166
Hidden-to-Visible Analysis.....	169
Advanced Training and Future Versions.....	172
Chapter 4: Greedy Training	173
Generative Sampling.....	175
Chapter 5: DEEP Operating Manual	183
Menu Options.....	183
File Menu Options.....	183
Test Menu Options	184
Display Menu Options.....	186
The "Read a database" Option.....	186
The "Read MNIST image" Option	186
The "Read MNIST labels" Option	187
The "Write activation file" Option.....	187
The "Clear all data" Option	187
Model Architecture.....	188
Database Inputs and Targets.....	189
RBM Training Params.....	189
Supervised Training Params	193
Train	196

TABLE OF CONTENTS

Test 198

Analyze..... 200

Receptive Field 201

Generative Sample..... 202

The DEEP.LOG File 206

Index..... 213