



3GE COLLECTION ON COMPUTER SCIENCE

DEEP LEARNING TECHNOLOGIES



3G E-LEARNING

3GE COLLECTION ON COMPUTER SCIENCE

DEEP LEARNING TECHNOLOGIES



3G E-LEARNING

**3GE COLLECTION ON COMPUTER SCIENCE:
DEEP LEARNING TECHNOLOGIES**



3G E-LEARNING

© 2019 3G E-learning LLC

90 Church Street

FL 1 #3514

New York, NY 10008

United States of America

www.3ge-learning.com

email: info@3ge-learning.com

Authored and Edited by 3G E-learning LLC, USA

ISBN: 978-1-98462-351-5

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, scanning or otherwise without prior written permission of the publisher.

This book contains information obtained from highly regarded resources. A wide variety of references are listed. Reasonable efforts have been made to publish reliable data and information, but the authors, editors, and the publisher cannot assume responsibility for the legality of all materials or the consequences of their use. The authors, editors, and the publisher have attempted to trace the copyright holders of all material in this publication and express regret to copyright holders if permission to publish has not been obtained. If any copyright material has not been acknowledged, let us know so we may rectify in any future reprint. Registered trademark of products or corporate names are used only for explanation and identification without intent to infringe.

Notice: Registered trademark of products or corporate names are used only for explanation and identification without intent of infringement. Case Studies and/or Images presented in the book are the proprietary information of the respective organizations, and have been used here specifically and only for educational purposes. Although care has been taken to check accuracy of formulas and procedures, the detailed methods should be tested further on a small scale before being adopted commercially.

For more information visit about 3G E-Learning LLC and its products, visit www.3ge-learning.com

Table of Contents

Preface

ix

Chapter 1	Machine Learning Basics	1
	Introduction.....	1
	1.1 Overview of Machine Learning (ML).....	2
	1.1.1 Purpose of Machine Learning.....	3
	1.1.2 How Machine Learning Works.....	4
	1.1.3 Examples of Machine Learning.....	5
	1.1.4 Machine Learning Applications.....	6
	1.1.5 The Future of Machine Learning.....	7
	1.2 Types of Learning.....	7
	1.2.1 Supervised Learning.....	8
	1.2.2 Unsupervised Learning.....	9
	1.2.3 Semi-supervised Learning.....	9
	1.2.4 Reinforcement Learning.....	9
	1.3 Types of Machine Learning Algorithms.....	10
	1.3.1 Supervised Learning Algorithm.....	10
	1.3.2 Unsupervised Learning Algorithm.....	10
	1.3.3 Reinforcement Learning Algorithm.....	10
	1.3.4 List of Common Machine Learning Algorithms.....	10
	Summary.....	51
	References.....	53
 Chapter 2	 Feedforward Neural Network	 55
	Introduction.....	55
	2.1 Process of Feedforward Neural Network.....	57
	2.1.1 Base of Recurrent Neural Network.....	59
	2.1.2 Feedforward Neural Network Structure.....	60
	2.1.3 Radial Basis Function Networks.....	64
	2.1.4 Normalization of Data.....	65
	2.1.5 Backpropagation Neural Network.....	67
	2.1.6 Single-Layer Perceptron.....	68
	2.1.7 Multi-Layer Perceptron.....	70
	2.2 Artificial Neural Network.....	71
	2.2.1 Behavior Artificial Neural Networks as Human Brain.....	73

2.2.2 Typical Structure of Artificial Neural Networks Model	75
2.2.3 Neurons' Methodology to determine Outputs	76
2.2.4 Common Activation Functions of Neuron	77
2.2.5 Preprocessing data for neural networks (The Data Scaling)	78
2.2.6 Classification of Learning Algorithms	79
2.2.7 Back Propagation Learning Algorithms of ANNs	80
2.2.8 The Mathematical steps of the computational procedures for the training algorithm into ANNs	83
2.2.9 Artificial Neural Network - Building Blocks	85
Summary	91
References	93

Chapter 3 Regularization for Deep Learning 95

Introduction	95
3.1 Regularization Techniques in Deep Learning	95
3.1.1 Regularization	96
3.1.2 Regularization help reduce Overfitting	97
3.1.3 Effective Capacity of Neural Networks	98
3.1.4 Implications	101
3.1.5 The Role of Regularization	102
3.1.6 Different Regularization Techniques in Deep Learning	105
3.2 Parameter Norm Penalties	106
3.2.1 L^2 Parameter Regularization	107
3.2.2 L_1 Regularization	110
3.3 Dropout	111
3.4 Early stopping	121
3.4.1 How early stopping acts as a regularizer	125
3.5 MNIST data with keras	127
Summary	137
References	138

Chapter 4 Optimization for Training Deep Models 139

Introduction	139
4.1 How Learning Differs from Pure Optimization	140
4.1.1 Empirical Risk Minimization	141
4.1.2 Surrogate Loss Functions and Early Stopping	142
4.1.3 Batch and Minibatch Algorithms	142
4.2 Challenges in Neural Network Optimization	147
4.2.1 Ill-Conditioning	147
4.2.2 Local Minima	148
4.2.3 Plateaus, Saddle Points and Other Flat Regions	150
4.2.4 Cliffs and Exploding Gradients	152

4.2.5 Long-Term Dependencies	153
4.2.6 Inexact Gradients	154
4.2.7 Poor Correspondence between Local and Global Structure	154
4.3 Basic Algorithms	156
4.3.1 Stochastic Gradient Descent	156
4.3.2 Momentum	158
4.3.3 Nesterov Momentum	161
4.4 Parameter Initialization Strategies	162
Summary	168
References	169
 Chapter 5 Convolutional Networks	 171
Introduction	171
5.1 Understanding Convolutional Neural Network	172
5.1.1 Defining a Convolutional Neural Network	175
5.1.2 Early CNN Development	181
5.2 Architecture of CNN	184
5.2.1 Back Propagation	185
5.2.2 Layers used to build ConvNets	186
5.2.3 ConvNet Architectures	202
Summary	207
References	208
 Chapter 6 Sequence Modeling: Recurrent and Recursive Nets	 211
Introduction	211
6.1 Unfolding Computational Graphs	213
6.2 Recurrent Neural Networks	216
6.2.1 Teacher Forcing and Networks with Output Recurrence	221
6.2.2 Computing the Gradient in a Recurrent Neural Network	223
6.2.3 Recurrent Networks as Directed Graphical Models	225
6.2.4 Modeling Sequences Conditioned on Context with RNNs	229
6.3 Bidirectional RNNs	231
6.4 Encoder-Decoder Sequence-to-Sequence Architectures	233
6.5 Deep Recurrent Networks	235
6.6 Recursive Neural Networks	236
Summary	239
References	240
 Chapter 7 Probabilistic Graphical Models	 241
Introduction	241
7.1 Basic Terminology and the Problem Setting	242

7.1.1 Bayesian Networks: Directed Graphical Models	243
7.1.2 Markov Networks: Undirected Graphical Models	246
7.2 Deep Learning and Graphical Models	254
7.2.1 Classic Neural Networks	255
7.2.2 Graphical Models	263
7.3 Deep Learning Architectures	265
7.3.1 Representation, Inference, Learning	266
7.4 Using Graphs to Describe Model Structure	268
7.4.1 Directed Models	268
7.4.2 Undirected Model	269
7.5 Deep Learning in Not Probabilistic Induction	271
Summary	275
References	277

Index

279

3GE COLLECTION ON COMPUTER SCIENCE

Deep Learning Technologies

Deep learning has emerged as a new area of machine learning research. It tries to mimic the human brain, which is capable of processing and learning from the complex input data and solving different kinds of complicated tasks well. It has been successfully pragmatic to several fields such as images, sounds, text and motion. Deep learning is a developing part of machine learning (ML) research. It contains multiple hidden layers of artificial neural networks. The deep learning methodology applies nonlinear transformations and model abstractions of high level in large databases. The recent advancements in deep learning architectures within different fields have already delivered significant contributions in artificial intelligence. The techniques developed from deep learning research have already been impacting the research of natural language process. Deep learning discovers intricate structure in large data sets by using the back propagation algorithm to indicate how a machine should change its internal parameters that are used to compute the representation in each layer from the representation in the previous layer. Deep convolutional nets have brought about breakthroughs in processing images, video, speech and audio, whereas recurrent nets have shone light on sequential data such as text and speech.

This book presents state of the art topics on deep learning, its applications and recent development in natural language processing. The book also presents how and in what major applications deep learning algorithms have been utilized.

Recently, machine learning and data mining have become the center of attention and the most popular topics among research community. These combined fields of study analyze multiple possibilities of characterization of databases. During the past several years, the deep learning techniques have already been impacting a wide range of machine learning and artificial intelligence. It is thought that moving machine learning closer to one of its original goals. Deep learning has becoming a new field of machine learning, and has gained extensive interests in different research area. It has shown some advantages over the traditional machine learning methods in some fields. Although deep learning works well in many machine learning tasks, it works equally poorly in some areas as the other learning methods. Besides most of the deep learning investigations are empirical, solid theoretical foundations of deep learning need to be established. Deep learning has been applied to natural language processing with some success.

Deep Learning Technologies provides an overview of general deep learning methodology and its applications to a variety of signal and information processing tasks. This book shows how faculty can help students develop skills in research, problem solving, critical thinking, and knowledge management by using web-based collaboration tools.



3G E-LEARNING

www.3ge-learning.com
email: info@3ge-learning.com

ISBN 978-1-98462-351-5



9 781984 623515