# Table of Contents

**About the Authors** ..... ix

**About the Technical Reviewers** ..... xi

**Acknowledgments** ..... xiii

**Introduction** ..... xv

**Chapter 1: Introduction to Machine Learning and Deep Learning** ..... 1

Defining Deep Learning ..... 1

A Brief History ..... 2

Advances in Related Fields ..... 8

Prerequisites ..... 9

The Approach Ahead ..... 9

Installing the Required Libraries ..... 10

The Concept of Machine Learning ..... 10

Binary Classification ..... 11

Regression ..... 13

Generalization ..... 14

Regularization ..... 22

Summary ..... 24

**Chapter 2: Introduction to PyTorch** ..... 27

Why Do We Need a Deep Learning Framework? ..... 27

What Is PyTorch? ..... 28

Why PyTorch? ..... 28

It All Starts with a Tensor ..... 29
Creating Tensors ..... 31
Tensor Munging Operations ..... 41
Mathematical Operations ..... 63
Element-Wise Mathematical Operations ..... 70
Trigonometric Operations in Tensors ..... 72
Comparison Operations for Tensors ..... 74
Linear Algebraic Operations ..... 76
Summary ..... 90

**Chapter 3: Feed-Forward Neural Networks ..... 93**

What Is a Neural Network? ..... 93
Unit ..... 94
The Overall Structure of a Neural Network ..... 95
Expressing a Neural Network in Vector Form ..... 97
Evaluating the Output of a Neural Network ..... 99
Training a Neural Network ..... 101
Deriving Cost Functions Using Maximum Likelihood ..... 102
Binary Cross-Entropy ..... 103
Cross-Entropy ..... 104
Squared Error ..... 106
Summary of Loss Functions ..... 106
Types of Activation Functions ..... 108
Linear Unit ..... 109
Sigmoid Activation ..... 109
Softmax Activation ..... 110
Rectified Linear Unit ..... 111
Hyperbolic Tangent ..... 112

Backpropagation ..... 113
Gradient Descent Variants ..... 114
Gradient-Based Optimization Techniques ..... 116
Practical Implementation with PyTorch ..... 119
Summary ..... 131

**Chapter 4: Automatic Differentiation in Deep Learning ..... 133**

Numerical Differentiation ..... 134
Symbolic Differentiation ..... 136
Automatic Differentiation Fundamentals ..... 137
Implementing Automatic Differentiation ..... 142
Summary ..... 144

**Chapter 5: Training Deep Leaning Models ..... 147**

Performance Metrics ..... 147
Classification Metrics ..... 148
Regression Metrics ..... 153
Data Procurement ..... 155
Splitting Data for Training, Validation, and Testing ..... 156
Establishing the Achievable Limit on the Error Rate ..... 157
Establishing the Baseline with Standard Choices ..... 158
Building an Automated, End-to-End Pipeline ..... 158
Orchestration for Visibility ..... 158
Analysis of Overfitting and Underfitting ..... 159
Hyperparameter Tuning ..... 160
Model Capacity ..... 161
Regularizing the Model ..... 163
Early Stopping ..... 164
Norm Penalties ..... 165

Dropout ..... 167
A Practical Implementation in PyTorch ..... 169
Interpreting the Business Outcomes for Deep Learning ..... 193
Summary ..... 195

**Chapter 6: Convolutional Neural Networks ..... 197**

Convolution Operation ..... 198
Pooling Operation ..... 206
Convolution-Detector-Pooling Building Block ..... 207
Stride ..... 211
Padding ..... 212
Batch Normalization ..... 212
Filter ..... 212
Filter Depth ..... 213
Number of Filters ..... 213
Summarizing key learnings from CNNs ..... 213
Implementing a basic CNN using PyTorch ..... 214
Implementing a larger CNN in PyTorch ..... 225
CNN Thumb Rules ..... 240
Summary ..... 242

**Chapter 7: Recurrent Neural Networks ..... 243**

Introduction to RNNs ..... 243
Training RNNs ..... 251
Bidirectional RNNs ..... 258
Vanishing and Exploding Gradients ..... 260
Gradient Clipping ..... 261

Long Short-Term Memory ..... 262
Practical Implementation ..... 265
Summary ..... 285

**Chapter 8: Recent Advances in Deep Learning ..... 287**

Going Beyond Classification in Computer Vision ..... 288
Object Detection ..... 288
Image Segmentation ..... 289
Pose Estimation ..... 291
Generative Computer Vision ..... 292
Natural Language Processing with Deep Learning ..... 294
Transformer Models ..... 295
Bidirectional Encoder Representations from Transformers ..... 295
GrokNet ..... 297
Additional Noteworthy Research ..... 299
Concluding Thoughts ..... 300

**Index ..... 301**

# Introduction

This book has been drafted with a unique approach. The second edition focuses on the practicality of the topics within deep learning that help the reader to embrace modern tools with the right mathematical foundations. The first edition focused on introducing a meaningful foundation for the subject, while limiting the depth of the practical implementations. While we explored a breadth of technical frameworks for deep learning (Theano, TensorFlow, Keras, and PyTorch), we limited the depth of the implementation details. The idea was to distill the mathematical foundations while focusing briefly on the practical tools used for implementation.

A lot has changed over the past three years. The deep learning fraternity is now stronger than ever, and the frameworks have evolved in size and adoption. Theano is now deprecated (ceased development); TensorFlow saw huge adoption in the industry and academia; and Keras became more popular among beginners and deep learning enthusiasts. However, PyTorch has emerged recently as a widely popular choice for academia as well as industry. The growing number of research publications that recently have used PyTorch over TensorFlow is a testament to its growth within deep learning.

On the same note, we felt the need to revise the book with a focus on engaging readers with hands-on exercises to aid a more meaningful understanding of the subject. In this book, we have struck the perfect balance, with mathematical foundations as well as hands-on exercises, to embrace practical implementation exclusively on PyTorch. Each exercise is supplemented with the required explanations of PyTorch's functionalities and required abstractions for programming complexities.

Part I serves as a brief introduction to machine learning, deep learning, and PyTorch. We explore the evolution of the field, from early rule-based systems to the present-day sophisticated algorithms, in an accelerated fashion.

Part II explores the essential deep learning building blocks. Chapter 3 introduces a simple feed-forward neural network. Incrementally and logically, we uncover the various building blocks that constitute a neural network and which can be reused in building any other network. Though foundational, Chapter 3 focuses on building a baby neural network with the required framework that helps to construct and train networks of all kinds and complexities. In Chapter 4, we explore the core idea that enabled the possibility of training large networks through backpropagation using automatic differentiation and chain rule. We explore PyTorch's Autograd module with a small example to understand how the solution works programmatically. In Chapter 5, we look at orchestrating all the building blocks discussed through so far, along with the performance metrics of deep learning models and the artifacts required to enable an improved means for training—i.e., regularization, hyperparameter tuning, overfitting, underfitting, and model capacity. Finally, we leverage all this content to develop a deep neural network for a real-life dataset using PyTorch. In this exercise, we also explore additional PyTorch constructs that help in the orchestration of various deep learning building blocks.

Part III covers three important topics within deep learning. Chapter 6 explores convolutional neural networks and introduces the field of computer vision. We explore the core topics within convolutional neural networks, including how they learn and how they are distinguished from other networks. We also leverage a few hands-on exercises—using a small MNIST dataset as well as the popular Cats and Dogs dataset—to study the practical implementation of a convolutional neural network. In Chapter 7, we study recurrent neural networks and enter the field of natural language processing. Similar to Chapter 6, we incrementally build an intuition

around the fundamentals and later explore practical exercises with real-life datasets. Chapter 8 concludes the book by looking at some of the recent trends within deep learning. This chapter is only a cursory introduction and does not include any implementation details. The objective is to highlight some advances in the research and the possible next steps for advanced topics.

Overall, we have put in great efforts to write a structured, concise, exercise-rich book that balances the coverage between the mathematical foundations and the practical implementation.

# Introduction

This book encapsulates our combined experience on what developers need to build resilient, scalable, and secured database applications. Azure SQL Database provides a solid foundation for database applications, and we believe this book will provide a solid foundation for the developers to build the applications for future.

**Chapter 1** starts off with the motivation for the book and sets the stage for developer mindset.

**Chapter 2** gets you kick-started with Azure SQL, how to provision and build your first database, and how to use the samples.

Once you have your database, you would like to connect to it and start querying it. **Chapter 3** gets you there. It also takes you through on configuring the right connection attributes and building the connection retry logic for resiliency.

**Chapter 4** takes you through the foundations of building the constructs and writing a database application.

**Chapter 5** builds on Chapter 4 and discusses more advanced concepts.

**Chapter 6** discusses practical use cases and best practices for using tables and indexes.

**Chapter 7** discusses more advanced concepts, such as consistency, scalability, and performance.

**Chapter 8** discusses the multi-model capabilities in Azure SQL that help you build modern applications.

**Chapter 9** discusses advanced data storage capabilities, such as columnstore indexes, memory-optimized tables, temporal tables, and so on.

Having written your application, you would like to monitor its performance and debug it.

**Chapter 10** takes you through that.

The journey of a developer isn't complete without the discussion on DevOps.

**Chapter 11** discusses DevOps technologies, tools, and processes for database applications involving Azure SQL database.