Introduction to Deep Learning Business Applications for Developers

From Conversational Bots in Customer Service to Medical Image Processing

Armando Vieira
Bernardete Ribeiro

Apress®

Introduction to Deep Learning Business Applications for Developers

From Conversational Bots in Customer Service to Medical Image Processing

Armando Vieira
Bernardete Ribeiro

Apress®

Introduction to Deep Learning Business Applications for Developers

Armando Vieira
Linköping, Sweden

Bernardete Ribeiro
Coimbra, Portugal

ISBN-13 (pbk): 978-1-4842-3452-5
ISBN-13 (electronic): 978-1-4842-3453-2
https://doi.org/10.1007/978-1-4842-3453-2

Library of Congress Control Number: 2018940443

Managing Director, Apress Media LLC: Welmoed Spahr
Acquisitions Editor: Celestin John
Development Editor: Matthew Moodie
Coordinating Editor: Divya Modi

Cover designed by eStudioCalamar

Cover image designed by Freepik (www.freepik.com)

Distributed to the book trade worldwide by Springer Science+Business Media New York, 233 Spring Street, 6th Floor, New York, NY 10013. Phone 1-800-SPRINGER, fax (201) 348-4505, e-mail orders-ny@springer-sbm.com, or visit www.springeronline.com. Apress Media, LLC is a California LLC and the sole member (owner) is Springer Science + Business Media Finance Inc (SSBM Finance Inc). SSBM Finance Inc is a **Delaware** corporation.

For information on translations, please e-mail rights@apress.com, or visit www.apress.com/rights-permissions.

Apress titles may be purchased in bulk for academic, corporate, or promotional use. eBook versions and licenses are also available for most titles. For more information, reference our Print and eBook Bulk Sales web page at www.apress.com/bulk-sales.

Any source code or other supplementary material referenced by the author in this book is available to readers on GitHub via the book's product page, located at www.apress.com/978-1-4842-3452-5. For more detailed information, please visit www.apress.com/source-code.

Printed on acid-free paper

To my family.

—Bernardete Ribeiro

Table of Contents

About the Authors

Armando Vieira earned his PhD in physics in 1997 from the University of Coimbra and started working in artificial neural networks soon after. He pioneered research on deep neural networks in 2003 and more recently worked as a senior data scientist consultant for several companies and startups, ranging from image processing, drug discovery, and credit scoring to risk analysis. He has been a speaker at many events related to artificial intelligence and business. He is the founder of Alea.ai. You can find more information at http://armando.lidinwise.com.

Bernardete Ribeiro is full professor at the University of Coimbra, Portugal, where she teaches programming, pattern recognition, business intelligence, and other topics. She holds a PhD and habilitation in informatics engineering at the University of Coimbra (CISUC). She is also the director of the Center of Informatics and Systems at CISUC. Her research interests are in the areas of machine learning, pattern recognition, financial engineering, text classification, and signal processing, as well as their applications in a broad range of fields. She has been the founder and director of the Laboratory of Artificial Neural Networks (LARN) for more than 20 years. Bernardete is the president of the Portuguese Association of Pattern Recognition (APRP) and member of the governing board of the International Association for Pattern Recognition (IAPR).

About the Technical Reviewer

Jojo Moolayil is an artificial intelligence, deep learning, machine learning, and decision science professional with more than five years of industrial experience. He is the author of *Smarter Decisions: The Intersection of IoT and Decision Science* and has worked with several industry leaders on high-impact and critical data science and machine learning projects across multiple verticals. He is currently associated with General Electric and lives in Bengaluru—the Silicon Valley of India.

He was born and raised in Pune, India, and graduated from the University of Pune with a major in information technology engineering. He started his career with Mu Sigma Inc., the world's largest pure-play analytics provider and has worked with the leaders of many Fortune 50 clients. One of the early enthusiasts to venture into IoT analytics, he converged his learnings from decision science to bring the problem-solving frameworks and his learnings from data and decision science to IoT analytics.

To cement his foundations in data science for industrial IoT and scale the impact of the problem-solving experiments, he joined a fast-growing IoT analytics startup called Flutura based in Bangalore and headquartered in the valley. After a short stint with Flutura, Jojo moved on to work with the leaders of industrial IoT, General Electric, in Bangalore, where he focuses on solving decision science problems for industrial IoT use cases. As a part of his role in GE, Jojo also focuses on developing data science and decision science products and platforms for industrial IoT.

In addition to authoring books on decision science and IoT, Jojo has been the technical reviewer for various books on machine learning, deep learning, and business analytics with Apress and Packt publications. He is an active data science tutor and maintains a blog at `www.jojomoolayil.com/web/blog/`. You can reach him at `https://www.linkedin.com/in/jojo62000`.

Acknowledgments

We would like to thank all those who have contributed to bringing this book to publication for their help, support, and input. In particular, we appreciate all the support and encouragement from ContextVision AB, namely, Martin Hedlund and Mikael Rousson, for the inspiring conversations during the preparation of the book.

We also want to thank the Center of Informatics and Systems of the University of Coimbra (CISUC) and to the Informatics Engineering Department and the Faculty of Science and Technologies at the University of Coimbra (UC) for the support and means provided while researching and writing this book.

Our thanks also to Noel Lopes who reviewed the technical aspects of the book related to multicore processing and to Benjamin Auffarth for the careful reading of the manuscript.

A special thanks and appreciation to our editors, Celestin John and Divya Modi, at Springer for their essential encouragement.

Lastly, thank you to our families and friends for their love and support.

Armando Vieira and Bernardete Ribeiro
Coimbra, Portugal
February 2018

Introduction

Deep learning has taken artificial intelligence by storm and has infiltrated almost every business application. Because almost all content and transactions are now being recorded in a digital format, a vast amount of data is available for exploration by machine learning algorithms. However, traditional machine learning techniques struggle to explore the intricate relationships presented in this so-called Big Data. This is particularly acute for unstructured data such as images, voice, and text.

Deep learning algorithms can cope with the challenges in analyzing this immense data flow because they have a very high learning capacity. Also, deep neural networks require little, if any, feature engineering and can be trained from end to end. Another advantage of the deep learning approach is that it relies on architectures that require minimal supervision (in other words, these architectures learn automatically from data and need little human intervention). These architectures are the so-called "unsupervised" of weakly supervised learning. Last, but not least, they can be trained as generative processes. Instead of mapping inputs to outputs, the algorithms learn how to generate both inputs and outputs from pure noise (i.e., generative adversarial networks). Imagine generating Van Gogh paintings, cars, or even human faces from a combination of a few hundred random numbers.

Google language translation services, Alexa voice recognition, and self-driving cars all run on deep learning algorithms. Other emergent areas are heavily dependent on deep learning, such as voice synthesis, drug discovery, and facial identification and recognition. Even creative areas, such as music, painting, and writing, are beginning to be disrupted by this technology. In fact, deep learning has the potential to create such a profound transformation in the economy that it will probably trigger one of the biggest revolutions that humanity has ever seen.

Thanks to the dissemination of free, and powerful, computational frameworks and APIs such as Keras and TensorFlow, cheap cloud services to run the models, and the easy availability of data, anyone can run deep learning models in their home in a matter of hours. This democratization helps to explain the explosion of interest in the topic and the many breakthroughs being presented in an open format on Arxiv and in specialized top conferences like NIPS.

Introduction to Deep Learning Business Applications for Developers explores various deep learning algorithms by neatly abstracting the math skills. It gives an overview of several topics focused on the business applications of deep learning in computer vision, natural language processing, reinforcement learning, and unsupervised deep learning. It is targeted to mid-level and senior-level professionals as well as entry-level professionals with a basic understanding of machine learning. You can expect to understand the tangible depth of business applications and view use-case examples regarding future developments in each domain.

The book gives a short survey of the state-of-the-art algorithms of the whole field of deep learning, but its main purpose is more practical: to explain and illustrate some of the important methods of deep learning used in several application areas and in particular the impact on business. This book is intended for those who want to understand what deep learning is and how it can be used to develop business applications, with the aim of practical and successful deployment. The book filters out any overwhelming statistics and algebra and provides you with methods and tips on how to make simple hands-on tools for your business model.

First it introduces the main deep learning architectures and gives a short historical background of them. This is followed by examples of deep learning that are most advantageous and that have promising futures over traditional machine learning algorithms. Along these lines, the book covers applications of recommendation systems and natural language processing, including recurrent neural networks capable of capturing the richness of exhibiting language translation models. The book finishes by looking at the applications of deep learning models for financial risk assessment, control and robotics, and image recognition. Throughout the text, you will read about key companies and startups adopting this technology in their products. You will also find useful links and some examples, tricks, and insights on how to train deep learning models with some hands-on code examples in Keras and Python.

PART I

Background and Fundamentals

CHAPTER 1

Introduction

This chapter will describe what the book is about, the book's goals and audience, why artificial intelligence (AI) is important, and how the topic will be tackled.

Teaching computers to learn from experience and make sense of the world is the goal of artificial intelligence. Although people do not understand fully how the brain is capable of this remarkable feat, it is generally accepted that AI should rely on weakly supervised generation of hierarchical abstract concepts of the world. The development of algorithms capable of learning with minimal supervision—like babies learn to make sense of the world by themselves—seems to be the key to creating truly general artificial intelligence (GAI) [GBC16].

Artificial intelligence is a relatively new area of research (it started in the 1950s) that has had some successes and many failures. The initial enthusiasm, which originated at the time of the first electronic computer, soon faded away with the realization that most problems that the brain solves in a blink of an eye are in fact very hard to solve by machines. These problems include locomotion in uncontrolled environments, language translation, and voice and image recognition. Despite many attempts, it also became clear that the traditional (rule-based and descriptive) approach to solving complex mathematical equations or even proving theorems was insufficient to solve the most basic situations that a 2-year-old toddler had no difficulty with, such as understanding basic language concepts. This fact led to the so-called long AI winter, where many

A. Vieira and B. Ribeiro, *Introduction to Deep Learning Business Applications for Developers*,
https://doi.org/10.1007/978-1-4842-3453-2_1

researchers simply gave up creating machines with human-level cognitive capabilities, despite some successes in between, such as the IBM machine Deep Blue that become the best chess player in the world or such as the application of neural networks for handwritten digit recognition in late 1980s.

AI is today one of the most exciting research fields with plenty of practical applications, including autonomous vehicles, drug discovery, robotics, language translation, and games. Challenges that seemed insurmountable just a decade ago have been solved—sometimes with superhuman accuracy—and are now present in products and ubiquitous applications. Examples include voice recognition, navigation systems, facial emotion detection, and even art creation, such as music and painting. For the first time, AI is leaving the research labs and materializing in products that could have emerged from science-fiction movies.

How did this revolution become possible in such a short period of time? What changed in recent years that puts us closer to the GAI dream? The answer is more a gradual improvement of algorithms and hardware than a single breakthrough. But certainly *deep neural networks*, commonly referred to as *deep learning* (DL), appears at the top of the list [J15].

1.1 Scope and Motivation

Advances in computational power, big data, and the Internet of Things are powering the major transformation in technology and are powering productivity across all industries.

Through examples in this book, you will explore concrete situations where DL is advantageous with respect to other traditional (shallow) machine learning algorithms, such as content-based recommendation algorithms and natural language processing. You'll learn about techniques such as Word2vec, skip-thought vectors, and Item2Vec. You will also consider recurrent neural networks trained with stacked long short-term memory (LSTM) units and sequence2sequence models for language translation with embeddings.

A key feature of DL algorithms is their capability to learn from large amounts of data with minimal supervision, contrary to shallow models that normally require less (labeled) data. In this book, you will explore some examples, such as video prediction and image segmentation, with fully convolutional neural networks (FCNNs) and residual neural networks (ResNets) that have achieved top performance in the ImageNet image recognition competition. You will explore the business implications of these image recognition techniques and some active startups in this very active field.

The implications of DL-supported AI in business is tremendous, shaking to the foundations many industries. It is perhaps the biggest transformative force since the Internet.

This book will present some applications of DL models for financial risk assessment (credit risk with deep belief networks and options optimizations with variational auto-encoder). You will briefly explore applications of DL to control and robotics and learn about the DeepQ learning algorithm (which was used to beat humans in the game Go) and actor-critic methods for reinforcement learning.

You will also explore a recent and powerful set of algorithms, named *generative adversarial neural networks* (GANs), including the dcGAN, the conditional GAN, and the pixel2pixel GAN. These are very efficient for tasks such as image translation, image colorization, and image completion.

You'll also learn about some key findings and implications in the business of DL and about key companies and startups adopting this technology. The book will cover some frameworks for training DL models, key methods, and tricks to fine-tune the models.

The book contains hands-on coding examples, in Keras using Python 3.6.

1.2 Challenges in the Deep Learning Field

Machine learning, and deep learning in particular, is rapidly expanding to almost all business areas. DL is the technology behind well-known applications for speech recognition, image processing, and natural language processing. But some challenges in deep learning remain.

To start with, deep learning algorithms require large data sets. For instance, speech recognition requires data from multiple dialects or demographics. Deep neural networks can have millions or even billion of parameters, and training can be a time-consuming process—sometimes weeks in a well-equipped machine.

Hyperparameter optimization (the size of the network, the architecture, the learning rate, etc.) can be a daunting task. DL also requires high-performance hardware for training, with a high-performance GPU and at least 12Gb of memory.

Finally, neural networks are essentially black boxes and are hard to interpret.

1.3 Target Audience

This book was written for academics, data scientists, data engineers, researchers, entrepreneurs, and business developers.

While reading this book, you will learn the following:

- What deep learning is and why it is so powerful
- What major algorithms are available to train DL models
- What the major breakthroughs are in terms of applying DL
- What implementations of DL libraries are available and how to run simple examples
- Major areas of the impact of DL in business and startups

The book introduces the fundamentals while giving some practical tips to cover the information needed for a hands-on project related to a business application. It also covers the most recent developments in DL from a pragmatic perspective. It cuts through the buzz and offers concrete examples of how to implement DL in your business application.

1.4 Plan and Organization

The book is divided into four parts. Part 1 contains the introduction and fundamental concepts about deep learning and the most important network architectures, from convolutional neural networks (CNNs) to LSTM networks.

Part 2 contains the core DL applications, in other words, image and video, natural language processing and speech, and reinforcement learning and robotics.

Part 3 explores other applications of DL, including recommender systems, conversational bots, fraud, and self-driving cars.

Finally, Part 4 covers the business impact of DL technology and new research and future opportunities.

The book is divided into 11 chapters. The material in the chapters is structured for easy understanding of the DL field. The book also includes many illustrations and code examples to clarify the concepts.

CHAPTER 2

Deep Learning: An Overview

Artificial neural networks are not new; they have been around for about 50 years and got some practical recognition after the mid-1980s with the introduction of a method (backpropagation) that allowed for the training of multiple-layer neural networks. However, the true birth of deep learning may be traced to the year 2006, when Geoffrey Hinton [GR06] presented an algorithm to efficiently train deep neural networks in an unsupervised way—in other words, data without labels. They were called *deep belief networks* (DBNs) and consisted of stacked restrictive Boltzmann machines (RBMs), with each one placed on the top of another. DBNs differ from previous networks since they are generative models capable of learning the statistical properties of data being presented without any supervision.

Inspired by the depth structure of the brain, deep learning architectures have revolutionized the approach to data analysis. Deep learning networks have won a large number of hard machine learning contests, from voice recognition [AAB+15] to image classification [AIG12] to natural language processing (NLP) [ZCSG16] to time-series prediction—sometimes by a large margin. Traditionally, AI has relied on heavily handcrafted features. For instance, to get decent results in image classification, several preprocessing techniques have to be applied, such as filters, edge detection, and so on. The beauty of DL is that most, if not

A. Vieira and B. Ribeiro, *Introduction to Deep Learning Business Applications for Developers*,
https://doi.org/10.1007/978-1-4842-3453-2_2

all, features can be learned automatically from the data—provided that enough (sometimes million) training data examples are available. Deep models have feature detector units at each layer (level) that gradually extract more sophisticated and invariant features from the original raw input signals. Lower layers aim to extract simple features that are then clumped into higher layers, which in turn detect more complex features. In contrast, shallow models (those with two layers such as neural networks [NNs] or support vector machine [SVMs]) present very few layers that map the original input features into a problem-specific feature space. Figure 2-1 shows the comparison between Deep Learning and Machine Learning (ML) models in terms of performance versus amount of data to build the models.

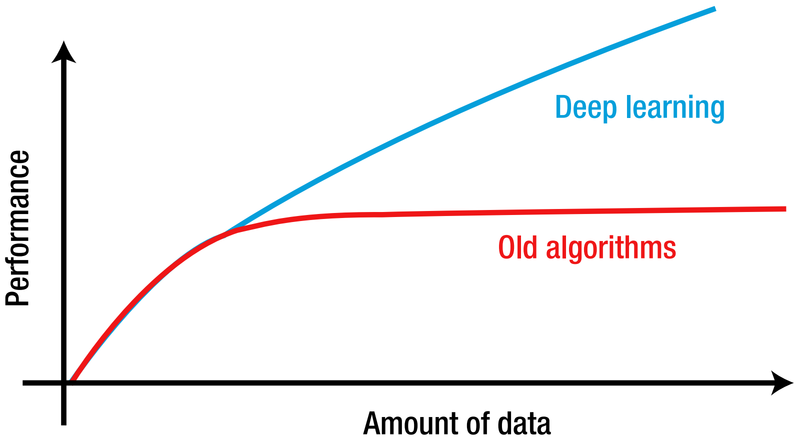

Figure 2-1. *Deep learning models have a high learning capacity*

Perfectly suited to do supervised as well as unsupervised learning in structured or unstructured data, deep neural architectures can be exponentially more efficient than shallow ones. Since each element of the architecture is learned using examples, the number of computational elements one can afford is limited only by the number of training samples—which can be of the order of billions. Deep models can be trained with hundreds of millions of weights and therefore tend to outperform shallow models such as SVMs. Moreover, theoretical results suggest that deep architectures are fundamental to learning the kind of complex functions that represent high-level abstractions (e.g., vision, language, semantics), characterized by many factors of variation that interact in nonlinear ways, making the learning process difficult.

2.1 From a Long Winter to a Blossoming Spring

Today it's difficult to find any AI-based technology that does not rely on deep learning. In fact, the implications of DL in the technological applications of AI will be so profound that we may be on the verge of the biggest technological revolution of all time.

One of the remarkable features of DL neural networks is their (almost) unlimited capacity to accommodate information from large quantities of data without overfitting—as long as strong regularizers are applied. DL is as much of a science as of an art, and while it's very common to train models with billions of parameters on millions of training examples, that is possible only by carefully selecting and fine-tuning the learning machine and sophisticated hardware. Figure 2-2 shows the trends in machine learning, pattern recognition and deep learning across the last decade/for more than one decade.

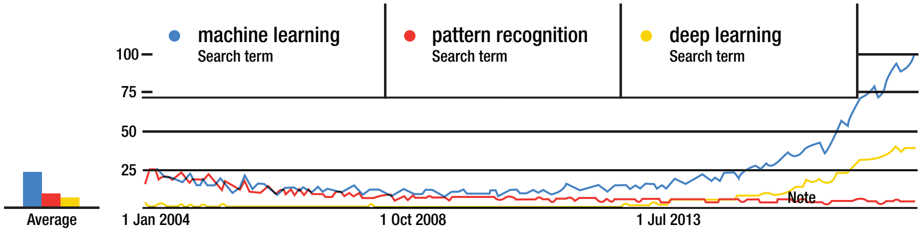

Figure 2-2. *Evolution of interest in deep learning (source: Google Trends)*

The following are the main characteristics that make a DNN unique:

- *High learning capacity*: Since DNNs have millions of parameters, they don't saturate easily. The more data you have, the more they learn.
- *No feature engineering required*: Learning can be performed from end to end—whether it's robotic control, language translation, or image recognition.
- *Abstraction representation*: DNNs are capable of generating abstract concepts from data.
- *High generative capability*: DNNs are much more than simple discriminative machines. They can generate unseen but plausible data based on latent representations.
- *Knowledge transfer*: This is one of the most remarkable properties—you can teach a machine in one large set of data such as images, music, or biomedical data

and transfer the learning to a similar problem where less of different types data is known. One of the most remarkable examples is a DNN that captures and replicates artistic styles.

- *Excellent unsupervised capabilities*: As long as you have lots of data, DNNs can learn hidden statistical representations without any labels required.
- *Multimodal learning*: DNNs can integrate seamlessly disparate sources of high-dimensional data, such as text, images, video, and audio, to solve hard problems like automatic video caption generation and visual questions and answers.
- They are relatively easy to compose and embed domain knowledge - or prioris - to handle uncertainty and constrain learning.

The following are the less appealing aspects of DNN models[1]:

- They are hard to interpret. Despite being able to extract latent features from the data, DNNs are black boxes that learn by associations and co-occurrences. They lack the transparency and interpretability of other methods, such as decision trees.
- They are only partially able to uncover complex causality relations or nested structural relationships, common in domains such as biology.

[1]Regarding these points, note that this is an active area of research, and many of these difficulties are being addressed. Some of them are partially solved, while others (such as lack of interpretability) probably never will be.

- They can be relatively complex and time-consuming to train, with many hyperparameters that require careful fine-tuning.
- They are sensitive to initialization and learning rate. It's easy for the networks to be unstable and not converge. This is particularly acute for recurrent neural networks and generative adversarial networks.
- A loss function has to be provided. Sometimes it is hard to find a good one.
- Knowledge may not be accumulated in an incremental way. For each new data set, the network has to be trained from scratch. This is also called the *knowledge persistence problem*.
- Knowledge transference is possible for certain models but not always obvious.
- DNNs can easily memorize the training data, if they have a huge capacity.
- Sometimes they can be easily fooled, for instance, confidently classifying noisy images.

2.2 Why Is DL Different?

Machine learning (ML) is a somewhat vague but hardly new area of research. In particular, pattern recognition, which is a small subfield of AI, can be summarized in one simple sentence: finding patterns in data. These patterns can be anything from historical cycles in the stock market to distinguishing images of cats from dogs. ML can also be described as the art of teaching machines how to make decisions.