LEARNING MADE EASY

HPE
Special Edition

Deep Learning

for dummies®
A Wiley Brand

Demystify deep learning

Unlock data insights for better business outcomes

Leverage deep learning solutions from HPE

Brought to you by

Hewlett Packard Enterprise

Scott D. Lowe

Hewlett Packard
Enterprise

HPE offers new solutions and capabilities to help organizations simplify the adoption of artificial intelligence (AI), with a focus on deep learning.

While you may be aware of the buzz around AI these days, with this book, we would like to make AI real for you with a focus on deep learning — regardless of where you are in the adoption process. We would like to be your partner while you understand what AI, deep learning, and machine learning can mean for you and your organization. This book walks you through the stages of adoption to optimizing solutions. It also provides information on hardware and software solutions from HPE that can meet your computing needs.

For more in-depth information on how HPE can help you in your AI journey please visit us at www.hpe.com/info/deep-learning.

Best regards,

Pankaj Goyal

Vice President, HPE AI Business
Hewlett Packard Enterprise

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Deep Learning

HPE Special Edition

by Scott D. Lowe

for dummies®
A Wiley Brand

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Deep Learning For Dummies®, HPE Special Edition

Published by
John Wiley & Sons, Inc.
111 River St.
Hoboken, NJ 07030-5774
www.wiley.com

Copyright © 2018 by John Wiley & Sons, Inc., Hoboken, New Jersey

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, except as permitted under Sections 107 or 108 of the 1976 United States Copyright Act, without the prior written permission of the Publisher. Requests to the Publisher for permission should be addressed to the Permissions Department, John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, (201) 748-6011, fax (201) 748-6008, or online at http://www.wiley.com/go/permissions.

Trademarks: Wiley, For Dummies, the Dummies Man logo, The Dummies Way, Dummies.com, Making Everything Easier, and related trade dress are trademarks or registered trademarks of John Wiley & Sons, Inc. and/or its affiliates in the United States and other countries, and may not be used without written permission. All other trademarks are the property of their respective owners. John Wiley & Sons, Inc., is not associated with any product or vendor mentioned in this book.

LIMIT OF LIABILITY/DISCLAIMER OF WARRANTY: THE PUBLISHER AND THE AUTHOR MAKE NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE ACCURACY OR COMPLETENESS OF THE CONTENTS OF THIS WORK AND SPECIFICALLY DISCLAIM ALL WARRANTIES, INCLUDING WITHOUT LIMITATION WARRANTIES OF FITNESS FOR A PARTICULAR PURPOSE. NO WARRANTY MAY BE CREATED OR EXTENDED BY SALES OR PROMOTIONAL MATERIALS. THE ADVICE AND STRATEGIES CONTAINED HEREIN MAY NOT BE SUITABLE FOR EVERY SITUATION. THIS WORK IS SOLD WITH THE UNDERSTANDING THAT THE PUBLISHER IS NOT ENGAGED IN RENDERING LEGAL, ACCOUNTING, OR OTHER PROFESSIONAL SERVICES. IF PROFESSIONAL ASSISTANCE IS REQUIRED, THE SERVICES OF A COMPETENT PROFESSIONAL PERSON SHOULD BE SOUGHT. NEITHER THE PUBLISHER NOR THE AUTHOR SHALL BE LIABLE FOR DAMAGES ARISING HEREFROM. THE FACT THAT AN ORGANIZATION OR WEBSITE IS REFERRED TO IN THIS WORK AS A CITATION AND/OR A POTENTIAL SOURCE OF FURTHER INFORMATION DOES NOT MEAN THAT THE AUTHOR OR THE PUBLISHER ENDORSES THE INFORMATION THE ORGANIZATION OR WEBSITE MAY PROVIDE OR RECOMMENDATIONS IT MAY MAKE. FURTHER, READERS SHOULD BE AWARE THAT INTERNET WEBSITES LISTED IN THIS WORK MAY HAVE CHANGED OR DISAPPEARED BETWEEN WHEN THIS WORK WAS WRITTEN AND WHEN IT IS READ.

ISBN 978-1-119-48458-5 (pbk); ISBN 978-1-119-48453-0 (ebk)

Manufactured in the United States of America

10 9 8 7 6 5 4 3 2 1

For general information on our other products and services, or how to create a custom *For Dummies* book for your business or organization, please contact our Business Development Department in the U.S. at 877-409-4177, contact info@dummies.biz, or visit www.wiley.com/go/custompub. For information about licensing the *For Dummies* brand for products or services, contact BrandedRights&Licenses@Wiley.com.

Publisher's Acknowledgments

Some of the people who helped bring this book to market include the following:

Development Editor: Elizabeth Kuball

Copy Editor: Elizabeth Kuball

Executive Editor: Katie Mohr

Editorial Manager: Rev Mengle

Business Development Representative: Karen Hattan

Production Editor: Tamilmani Varadharaj

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Table of Contents

INTRODUCTION 1
About This Book 1
Foolish Assumptions 1
Icons Used in This Book 2
Beyond the Book 2

CHAPTER 1: **Artificial Intelligence: All Buzz or the Key to Your Reality** 3
Defining Artificial Intelligence 4
Defining Machine Learning 5
Defining Deep Learning 8
Seeing What's Driving Artificial Intelligence and Deep Learning 10
Making Artificial Intelligence and Deep Learning Practical and Attainable 11

CHAPTER 2: **Considering IT Challenges with Deep Learning Adoption** 13
Suboptimal Infrastructure 13
Inability to Scale and Integrate 14
Limited Knowledge, Resources, and Capabilities 14

CHAPTER 3: **Demystifying Deep Learning with HPE** 15
Rapidly Developing Deep Learning Models 15
Scaling and Integrating Artificial Intelligence and Deep Learning Capabilities 17
Optimizing the Environment to Unlock Competitive Advantage 19

CHAPTER 4: **Developing Deep Learning Models** 21
Selecting Ideal Technologies and Systems 21
Training Your Teams Using the Deep Learning Institute 24
Understanding Sample Artificial Intelligence Solutions 24

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

CHAPTER 5: **Delivering Attractive Returns and Results** 27
Improving Deployment Effectiveness 27
Accelerating Analytics to Improve Business Operations 28
Integrating Artificial Intelligence and Deep Learning in Your Environment 29
Seeing How Artificial Intelligence and Deep Learning Impact Industries 30

CHAPTER 6: **Enhancing Your Competitive Advantage** 33
HPE Hardware Platforms 33
How Accelerated Graphics Processing Unit Compute Impacts Learning Speed 39

CHAPTER 7: **Ten (Or So) Ways HPE Can Help You Get Started** 41

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Introduction

Self-driving cars! Automation! Robots taking over jobs! These are just a few of the things that the media covers when the topic of artificial intelligence (AI) comes up. Of course, there is far more to the story, and there are far more everyday ways that the burgeoning fields of AI, machine learning, and deep learning can transform your business, your career, and your life.

As the world becomes more intertwined with technology, digital transformation is rapidly reshaping every industry. Some industries have already transformed, while others — like retail banking, government, and manufacturing — are under tremendous pressure to innovate just to survive. AI plays a huge role in the era of digital transformation.

Understanding the full scope of this ultra-hyped field will soon be a make-or-break component for many careers and businesses. Companies that fail to stay current with the AI wave will find themselves left in the dust as competitors surge past them.

About This Book

AI and its components have the potential to reshape society as we know it. In this book, you learn about all the areas that comprise the AI field and gain an understanding for where each plays a significant role. I also bust some myths about AI and provide you with comprehensive information about how HPE's products and services can help companies like yours prepare for this major change in the way that business is done.

Foolish Assumptions

In writing this book, I assumed that you have at least a basic understanding of data center computing and virtualization. The general audience for this book is anyone in IT who may want to learn more about data center architectures. You may be part of your organization's technical staff, or you may be managerial or executive staff. Either way, this book is for you!

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Icons Used in This Book

Throughout this book, you find a number of icons intended to help you better understand and remember key concepts.

The Remember icon points out anything that's important enough to commit to your long-term memory.

Anything marked with the Tip icon helps you save time or money or just generally makes your life a little easier.

I don't dive super deep into technical stuff in this book, but there are some technical elements. I've marked them with the Technical Stuff icon.

Beyond the Book

For more information on Deep Learning with HPE, visit `www.hpe.com/info/deep-learning`.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

IN THIS CHAPTER

- Defining artificial intelligence, machine learning, and deep learning
- Understanding why artificial intelligence and deep learning are popular
- Adopting artificial intelligence

Chapter 1
Artificial Intelligence: All Buzz or the Key to Your Reality

So, what's all the buzz around artificial intelligence (AI)? AI has been showing potential for years, but interest seems to have absolutely exploded recently. Analyst firm IDC predicts that global spending on cognitive and AI solutions will increase at a 54.4 percent compound annual growth rate (CAGR) over the next few years, exceeding $46 billion by 2020. That's a lot of money.

Clearly, no business can — or should — avoid the AI reality. To run a business effectively, you need information and intelligence. One of the key reasons that AI has attracted a high level of corporate interest is because it can discover patterns in big data (and much, much more).

And that interest is growing. In fact, you interact with AI on a daily basis. Every time you ask Apple's Siri to call home or you order Amazon's Alexa to tell you the current weather, you're directly interacting with AI-based systems, at a personal level. These personal AI systems are good for relatively simple, repeatable tasks, and they're becoming far more common.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

For you and your business, though, AI may not yet be a significant reality. You may be eager to realize the benefits of AI, including saving your company time and money. For example, you may want to deploy enhancements to your customer service processes. Perhaps an AI-driven chat-bot answering a customer query in just seconds is cheaper than a human answering in minutes, saving your company money while also making the customer very happy. If you use a service such as Uber, you're buying a service from a company that uses machine learning to create efficient pickup and routing services. Pinterest uses machine learning to combat spam and to improve monetization. These are things that people use every day.

One of the biggest challenges people face when starting a journey down the AI road is understanding what, exactly, is meant by terms like *artificial intelligence, machine learning,* and *deep learning.* The following sections help you make sense of the madness.

Defining Artificial Intelligence

TIP

AI is the capability of a machine to imitate intelligent human behavior. AI is not a new concept. In fact, since the beginning of the information age in the 1950s, we have experienced AI in many forms. For example, early on, computer programs allowed systems to play games like checkers against human opponents. Today, Libratus, an AI, defeated four top human specialists in a 120,000-hand competition of heads-up no-limit Texas hold'em, the leading benchmark and long-standing challenge problem in imperfect-information game solving.

REMEMBER

The term *artificial intelligence* is used to describe the broad field and, itself, can have many different meanings and uses.

The original goal around AI was to create machines with the characteristics of human intelligence. These thinking, reasoning machines would be able to solve complex problems across the spectrum of human capabilities, with access to replicas of all our senses as inputs.

Today, we are realizing the promise of AI using machine and deep learning techniques in our everyday lives, with examples like Amazon's Alexa, Google Maps, and Apple's Siri.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Defining Machine Learning

Simple mimicking of human behavior isn't really all that hard. Going beyond that requires a bit more effort, though, and that's where machine learning comes in. Machine learning is a subset of artificial intelligence that gives computers the ability to learn without being explicitly and rigidly programmed. The algorithms that exist in a machine learning environment enable systems that can go beyond traditional strict programming processes.

How does this work in action, though? You've probably used an automated telephone system that takes your voice response so that you can be connected to the correct party when you call a big company. And if you're like me, you get frustrated at just how easily such systems can be confused! They're rigidly programmed to respond to a strict voice pattern.

But what if that automated system could learn based on its own experiences? As each caller dials in, this automated attendant, while fulfilling user requests, keeps track of where it made mistakes and everything that users did, and then self-adjusts in subsequent interactions. Or perhaps it looks for common patterns and identifies ways to streamline the process.

This is the realm of machine learning and there are a couple of different ways that it can be accomplished.

Computers are taught, or self-learn, how to recognize things — for example, the following:

- Animals in pictures
- The lanes on a freeway
- A computer hacker's digital fingerprints
- Predict hard drive failures
- Lung cancer from 3D image

There are innumerable uses for machine learning when the computer has figured out how to do the recognition, also referred to as the *training*. This initial training can be very hard work, both for the humans who prepare the training data and for the computer that's trying to create the models based on this training data.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

AI BRINGS LIFE TO COMPUTERS

When you step back and think about it, the machine learning process is really a lot like our own development process, even if we don't remember it actually taking place! As the machine gains experience, it's able to more efficiently learn new capabilities, just like human beings. The more life experience you have, the more of a baseline you have from which to work.

But a ton of research is going into making the machine learning training phase easier. For example, right now, it takes millions of examples to train a machine learning system to identify what a bird looks like. We humans can do it much, much more quickly, at least for now. As you train the machine, it becomes more efficient and, in fact, learns faster with iteration.

After the machine learning has completed its initial training, it uses ongoing data to continually update its models — it's always learning. For example, a system that predicts the failure of water pumps may be successful in those predictions 92 percent of the time. But in 8 percent of cases, it either misses a failure or predicts a failure when none is imminent. These are considered either false negatives or false positives. For these failures, the system begins to ask itself some questions, including, "What is it about my learning model that causes these false predictions, these anomalies?" Hopefully the system can learn and improve. Ideally, this learning can happen independently, but when it can't, people who understand machine learning can give it a boost.

There are different methods by which machines can learn. I cover these methods in the following sections.

Supervised learning

In supervised learning, machines are programmed with training data sets, which ensure that the output is of a certain value or quality. Supervised learning is, today, the most common way that machines are trained.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

If you didn't think there were already enough challenges to explore in the world of AI, supervised learning is itself broken down into two major categories:

- **Classification:** In classification-based supervised learning, machines are taught to classify or categorize certain inputs based on their programming. The outcome is often a state or a characteristic. For example, if you're teaching a machine how to identify birds, classification could be "This is a bird" or "This is not a bird." It's possible for more complex classification processes to be implemented using more advanced algorithms.
- **Regression:** Regression learning allows a machine to estimate values, such as future 401(k) earnings, based on the input data you provide the machine. Rather than a simple categorization, you get value-based results. The more data you feed the machine, the more it can learn and the more accurate it can be with its results.

Unsupervised learning

There's a lot more to automated security than just monitoring for known breach vectors, and that means that machine learning tools sometimes need to act independently of their human supervisor. Attackers are always adjusting their methods because they know that simple systems can't keep a watchful eye on everything all the time.

REMEMBER

Cyber-attacks are similar to the attacks on the human immune system. All kinds of viruses attacking our bodies and morphing all the time. And just like with the human immune system, the defenders must also morph, too.

In simple terms, unsupervised learning starts with an algorithm and a data set. Supervised learning also includes an output variable of some kind, but unsupervised learning relies on the machine to discover its own meaning in the data you provide it.

Not every tough problem is the same. There are a couple of categories in the world of unsupervised AI learning as well:

- **Clustering:** In a clustering problem, you rely on the machine to identify common characteristics in your underlying data.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

For example, you may cluster your client database based on which services they've used over the past year. From a business perspective, you may use this information to target those customers for new services.

- **Association:** Association-based machines look for ways to discover data patterns that may not be that obvious. For example, you may discover that most of the clients that purchase one service also purchase a second service. With this information, you may discover that bundling these services increases your revenue as more customers buy the bundle. That's just a simple way that AI can help improve your business.

Defining Deep Learning

Machine learning is certainly powerful, but deep learning takes it to a whole new level. Deep learning is an evolution of machine learning that creates increasingly complex hierarchical models intended to mimic thought processes in the human brain more than simple machine learning models.

As such, deep learning is a generation beyond machine learning, while still being a subset of both AI and an evolution of machine learning. Like so many technologies people have never heard of, deep learning may already have an impact on your life.

People are delighted to discover that Netflix knows what they want to watch. They gaze in amazement as Amazon displays other products that they may want to buy.

Although magical, this isn't the work of the black arts. These kinds of recommendation engines are driven by deep learning systems that can act on vast quantities of data to help improve the quality of your life.

Machine learning is the mainstream today, but deep learning has the most transformative potential to solve real-world problems that go far beyond simple entertainment. Figure 1-1 sums up the AI landscape.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

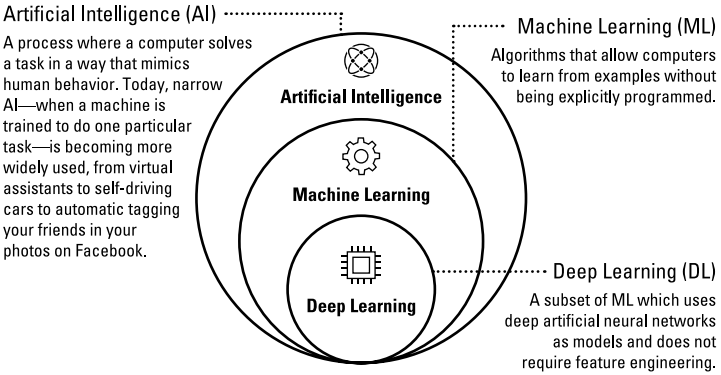


FIGURE 1-1: A visual comparison of AI technologies.

To summarize, with rule-based AI and machine learning, a data scientist has primary responsibility for deciding which rules and data set features will be included in models, which drives how those models operate. With deep learning, data scientists submit raw data into an algorithm. The system analyzes that data and, based on what it knows and what it can infer from the new data you've presented, it makes a prediction.

ARTIFICIAL INTELLIGENCE AND HIGH-PERFORMANCE COMPUTING

High-performance computing (HPC) clusters have been around for ages, but they've found new customers thanks to the insatiable thirst of AI. HPC and AI are like peanut butter and chocolate. Separate, they're both really tasty, but together, you get a special treat that will delight your senses!

Traditionally, HPC systems have addressed complex data problems. With the ability to support large data sets, in-memory databases, and complex models, HPC solutions have the horsepower necessary to help organizations gain insight into their data.

As the foundation for cognitive computing, organizations seeking to harness AI to help meet their goals need advanced technology solutions, such as HPC. These solutions enable AI applications to execute increasingly difficult tasks and forecast evolving trends, equipping them to solve some of the world's biggest scientific, engineering, and technological problems.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Seeing What's Driving Artificial Intelligence and Deep Learning

No matter how incredible a technology appears to be, it won't gain any traction if there aren't clear business outcomes that can be attained. Deep learning is no different. Thanks to the potential to surpass the competition, though, there are a number of reasons that deep learning is going mainstream.

Gaining insight from increasing quantities of data

Data is the new currency. Whoever has the data will win the day. Deep learning crunches data in ways that simple rules-based AI systems can only dream. This insight can be directly responsible for increased revenue generation or for improving cancer detection or for averting crop diseases in agriculture, and for so much more.

Understanding the business potential of deep learning

As a bank, how do you know — *really* know — that the customer that you're prepared to loan money to will pay it back? You have his credit score to work with, which provides you with a good snapshot of his credit profile, but what if you could do more? What if you had a deep learning–backed AI system that could scour the applicant's history and discover that he's a fraudster? You might save your company a lot of money!

REMEMBER

Perhaps even more significant is the reverse. Maybe an applicant with a terrible credit score is actually a decent risk. A simple credit score doesn't tell the whole story, but with an automated system sitting next to you, crunching data in an effort to provide you with a more complete picture, you may decide that the applicant is worth the risk and provide the loan. At the same time, your system may discover that the applicant is suitable for other products you make available. With the right data and the right insights, you have the potential to dramatically enhance your organization's bottom line.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

In agriculture, how do you leverage data to detect crop diseases so that you can remove them before they create widespread damage? Deep learning algorithms in agriculture are already doing exactly this, resulting in increased crop yields.

In a manufacturing environment, deep learning is being put to use to enable voice-based picking of warehouse parts. Voice is a far more efficient and natural way to interact with inventory control systems, so these types of deep learning systems can reduce costs and improve accuracy.

Making Artificial Intelligence and Deep Learning Practical and Attainable

AI and deep learning are technologies that may seem like science fiction to you. We all want a Commander Data by our sides, but instead we get a rules engine to configure. The right mix of software, hardware, and services is critical to leveraging the nascent capabilities of deep learning and modern AI. The best part is that all the ingredients for an incredible deep learning solution are available to you today, and adoption may not be the pipe dream you thought it would be.

See how HPE is helping customers navigate their AI journey through the development of a three-phase process: Explore, Experiment, and Expand.

Explore

REMEMBER

If you haven't discovered the real potential of using AI, now is the time to jump in and explore. During this exploratory phase, you need to determine what's possible and how it may impact your business. This is also where you can test the limits of the AI hype cycle to see what's real and what's not. It's also a stage at which engaging a trusted partner such as HPE can pay huge dividends. As an advisor, HPE can help you understand your options and, with you, create an adoption plan that considers the infrastructure you need to get started.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Experiment

Now, play! What are the limits for what you've put into place? How do you make it self-aware enough that it identifies and assists with the critical aspects of your business and makes you more competitive?

HPE can be your partner in your AI journey. Through the company's Centers of Excellence, you can accelerate your deep learning applications and realize better return on investment (ROI) from your deep learning deployments in the near term.

Expand

After you've completed your experiments, try it in production, where you can begin to realize the ROI that makes deep learning such a compelling technology. As you discover more services or opportunities that can benefit from AI, you can expand your deep learning footprint to encompass these new use cases. Here, too, HPE can help you implement and grow your environment to support even the most complex services you may dream up.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

IN THIS CHAPTER

- Seeing what's holding you back from realizing your machine-driven future
- Getting help with implementing deep learning

Chapter 2
Considering IT Challenges with Deep Learning Adoption

As you consider venturing down the deep learning road, you probably anticipate running into road blocks along the way. In this chapter, I fill you in on some of the key IT challenges that can prevent companies from realizing the full potential of implementing deep learning solutions.

Suboptimal Infrastructure

Unfortunately, traditional infrastructure doesn't meet the extreme compute requirements that are needed for deep learning capabilities.

Unlike traditional applications, deep learning solutions require a lot of parallelism, performance processors, dense storage, a lot of power, and a lot of cooling. Traditional infrastructure may need those things as well, but one of the main ways in which deep

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

learning infrastructure is different from traditional infrastructure is that deep learning relies on high-performance accelerators such as graphics processing units (GPUs).

Inability to Scale and Integrate

Let's suppose you get past the infrastructure hurdle. After you've developed a deep learning model and conducted your very first proof of concept (POC), you want to get it into production as quickly as possible. You want to get value out of your deep learning investment. You want to move from a POC to production, but you may find yourself struggling with your ability to scale and integrate.

You may lack the resources and expertise necessary to move out of POC into a production environment. Many projects have these kinds of challenges and require a boost to get beyond that, to a point at which the business is benefitting.

REMEMBER

An AI project that is stalled won't yield any results. You may be stalled going into production. You may be stalled in a learning model. Your AI may be stalled at its current level of development due to a lack of infrastructure resources. HPE is your proven partner, here to get you out of being stalled and on your way to faster insights and competitive advantage.

Limited Knowledge, Resources, and Capabilities

As an emerging field, AI skills can be difficult to procure. To get the right skills in a direct hire can be expensive and challenging. To expand your AI footprint, you may need help optimizing your environment, leveraging the latest in technologies, and you may need to partner with a vendor to help. Fortunately, HPE has the capability to help you leverage the latest technologies to gain competitive advantage from your deep learning deployment. From the company's AI Innovation Centers to HPE Pointnext deployment services, HPE can bring skills and experience to bear and, when coupled with the right infrastructure, can help customers bring AI to their lines of business.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

IN THIS CHAPTER

- Streamlining artificial intelligence training models
- Discovering resources that can help you on your deep learning journey

Chapter 3
Demystifying Deep Learning with HPE

Deep learning is a complex topic. Many people simply don't know where to begin with the technology, let alone understand what they can ultimately do with it. That's where HPE comes in. Through a combination of products, software, and services, HPE can help companies decode deep learning to help them benefit.

In this chapter, I walk you through how HPE can help.

Rapidly Developing Deep Learning Models

Developing models can be challenging. Sometimes, you just need help to get you past this initial level. That's where HPE comes in.

Leveraging deep learning IT expertise

Deep learning workloads have unique requirements, particularly when it comes to training modules. Choosing an ideal technology platform isn't simple.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

HPE can help IT departments create the optimal foundation for deep learning with rich expertise and tools to select the right combination of central processing unit (CPU) and performance accelerators, like graphics processing unit (GPU) technologies, and system platforms to meet their unique requirements. Many IT departments, data scientists, and line of business decision makers lack a good understanding of deep learning frameworks and infrastructure requirements.

A deep learning framework provides the building blocks for the creation and training of deep learning systems. A number of such frameworks are available, including the Microsoft Cognitive Toolkit, TensorFlow, Caffe, and many more. Each framework carries with it a different set of operations. With the number of frameworks available, choosing the right one for your needs can be confusing. Some of the frameworks are easier to use and learn than others as well. To learn more about some of the more common frameworks, go to `http://hpe.to/6004DdXVK`.

HPE can help to train data scientists and IT staff on the various deep learning frameworks and the supporting infrastructure requirements with hands-on or online training.

For those data scientists who are looking to gain quick access to an ideal platform to develop deep learning models, HPE has several options to enable rapidly developed deep learning models.

REMEMBER

Through a mix of comprehensive hardware, software, services, and other resources, HPE makes it possible for any organization to harness the power of AI.

Providing solutions

HPE has created one such integrated solution through the combination of the HPE Apollo 6500 Gen 10 System and Bright Computing software. This is an ideal platform to develop deep learning training models. It comes with the maximum GPU capacity and the deep learning frameworks that data scientists and application developers can use to get started. It also provides significant flexibility and choice in terms of both hardware and software.

HPE has found that many IT departments are under pressure from CIOs and CEOs to embark on AI initiatives.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

If you'd like to take a bigger step and use existing software applications, HPE can help you build solutions that can help with, among other things, fraud detection in e-commerce applications. This is just one example of solutions that can be deployed.

Scaling and Integrating Artificial Intelligence and Deep Learning Capabilities

As you begin your deep learning expedition, you'll probably find that you're able to create some simple models on your own. But as you move these models out of development and into production, you may find challenges in scaling the infrastructure and integrating the massive data volumes to deliver real returns to the business.

Providing proven blueprints

In your journey to overcome these challenges, HPE has created an entire practice around helping organizations across the globe implement deep learning solutions.

Reference architectures

If you're interested in building the infrastructure for a specific use case, such as image classification, HPE provides reference configurations with hardware platforms and sizing guides that will help you to successfully scale production deployments.

REMEMBER

But hardware alone does not a functional AI cluster make. Managing such clusters requires specialized tools. To this end, HPE makes available cluster management software to scale management for large deployments.

AI Innovation Centers

HPE makes available two key resources to help customers in their AI journeys. The first, the HPE AI Innovation Center, is designed for longer-term research projects. The Innovation Center serves as a platform for research collaboration between universities, enterprises on the cutting edge of AI research, and HPE researchers.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

The centers, located in Houston, Palo Alto, and Grenoble, give researchers for academia and enterprises access to infrastructure and tools to continue research initiatives.

The second resource, HPE Centers of Excellence (CoE), are designed to assist IT departments and data scientists who are looking to accelerate their deep learning applications and realize better ROI from their deep learning deployments in the near term. The HPE CoE offers select customers access to the latest technology and expertise, including the latest NVIDIA GPUs on HPE systems. The current CoE are spread across five locations including Houston; Palo Alto; Tokyo; Bangalore, India; and Grenoble, France.

TIP

Here are just some of the ways that the AI Innovations labs and Centers of Excellence can help you get started with your deep learning projects:

- **Driving technology development:** Integrate technologies required for deep learning and AI more quickly.
- **Demonstrating insightful benchmarking and testing:** Understand how solutions perform before fully investing.
- **Developing reference architectures and configurations:** Identify solutions specific to your requirements.
- **Validating and developing software:** Prove that the software performs as expected.

Both internal HPE researchers and leading researchers in academia working on AI projects need a stable, AI sandbox. The AI sandbox is a testing environment that isolates untested code changes and outright experimentation from the production environment in the context of software development. They need access to significant but not excessive hardware with comprehensive AI environment while avoiding the need to re-provision for each use. HPE has resources that can help.

Deploying, integrating, and supporting a deep learning environment

Most organizations can't just go it alone with AI. To assist, HPE provides comprehensive services for deployment, data integration,

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

and support. Specifically, HPE Pointnext provides a wide array of services around deep learning and AI, including the following:

- **Advisory services:** With AI experts, HPE Pointnext can point you in the right direction to help realize quicker time-to-value from your AI investment.
- **Professional and operational services:** Are you looking for on-time implementation and on-budget execution for your AI project? HPE Pointnext provides you with full-spectrum support to ensure that both happen.

Bringing flexible, on-demand capacity to the deep learning game

With the HPE GreenLake Flex Capacity service, HPE brings increased financial flexibility and control to your on-premises AI infrastructure deployment, which carries the following characteristics:

- HPE owns and supports your choice of on-premises equipment.
- Pay for what you use based on metered usage, and access self-service reporting to forecast future demand.
- An on-site buffer allows you to quickly scale up capacity when needed and avoid spending on unused capacity.

Optimizing the Environment to Unlock Competitive Advantage

After you successfully develop deep learning models and get them into production, there is an opportunity to leverage the latest technology innovations to accelerate performance and optimize environments that can deliver significant competitive advantage. HPE products and services are well positioned to make this happen.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Surpassing your competition with global centers of excellence

HPE's Centers of Excellence help you leverage the latest technologies, modernize your application code, and demonstrate the outcomes. In these centers, you can also perform proof-of-concept activities and benchmarking. These centers of excellence are customer facing and meant for larger customer engagements where they can get access to expertise and tools.

Leveraging next-gen technology integration

HPE doesn't do all this alone. The company partners with top-tier solution providers delivering core technologies that drive the solution. HPE tests and verifies these platforms to ensure that you're able to gain the best in terms of next-gen technology expertise.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

IN THIS CHAPTER

- Choosing the right technologies and systems for you
- Finding out how the Deep Learning Institute can help
- Looking at sample artificial intelligence solutions

Chapter 4
Developing Deep Learning Models

Lots of resources are available to you as you begin developing your own deep learning models. This chapter shows you where to find the resources you need.

Selecting Ideal Technologies and Systems

TIP

The success of your deep learning initiative is dependent on the robustness of the underlying technology. To help you get started creating models right away, HPE has a number of resources you can use.

Understanding HPE's Deep Learning Cookbook

The HPE Deep Learning Cookbook is a comprehensive set of tools to guide the choice of the best hardware and software environment

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

for different deep learning workloads. It consists of three main components:

- **Deep Learning Benchmarking Suite:** A benchmarking tool to collect performance measurements on various hardware and software configurations in a unified way
- **Deep Learning Performance Guide:** A web-based tool to analyze performance data collected with the Benchmarking Suite, and to predict performance of untested hardware and software configurations with analytical performance models
- **Reference Designs:** Default hardware and software recipes for select deep learning workloads, created based on extensive benchmarking and performance analysis for these workloads

Today's benchmarking suite supports a wide range of the most popular deep learning frameworks (BVLC Caffe, NVIDIA Caffe, Intel Caffe, Caffe2, TensorFlow, MXNet, PyTorch and NVIDIA inference engine TensorRT) and 18 popular deep models as reference workloads (see the full list at `http://hpe.to/6008DdXzk`). Using the benchmarking suite, HPE has collected massive amounts of performance data on multiple HPE hardware systems, including HPE Apollo 6500, HPE Apollo 6000, HPE Apollo 2000, HPE Edgeline 4000 and 1000, and HPE ProLiant DL380. This performance data is publicly available through HPE's Deep Learning Performance Guide, a second component of the Cookbook.

TIP

So, what exactly is the Cookbook? It's a combination of an automated benchmarking tool (to measure performance of different hardware and software stacks), an analysis tool, and a set of default hardware configurations for selected deep learning workloads (Reference Designs). You use the benchmarking tool to measure and validate the overall performance of your environment. Then you use the web-based analysis tool to leverage a massive collection of performance results for different deep learning workloads on different hardware and software stacks and analytical performance models.

There are key use cases that the Cookbook can help support — it's not just for newbies. For example, maybe you started at a small scale and you've seen initial success. Now you want to scale up

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

the infrastructure. Scaling up is an altogether different problem. HPE's Deep Learning Cookbook helps you understand performance bottlenecks and design a rock-solid scale-up, scale-out architecture.

Or, maybe you started in the public cloud, but you want to bring your AI deployment back on-premises. You may have chosen the public cloud to get started and seen some initial success. Now the line of business or the CIO wants to bring back the algorithms on-premises to get the best cost, security, performance, and scalability. The Deep Learning Cookbook can help you design the optimal architecture to bring back the model on-premises.

Getting expert advice to jumpstart your initiative

Let's dive deeper into the HPE Deep Learning Cookbook in action, helping customers meet their goals.

TIP

Trying to get started, but don't know how?

If you're like most people, you're fascinated by AI. You want to get started, but you may lack the expertise and resources to build the model and the right infrastructure to run it. Data scientists want to focus on building the *algorithm*, rather than building the hardware/software infrastructure to run it. The HPE Deep Learning Cookbook recommends easy get-started configurations. It also recommends which algorithms to use for which applications.

Using the HPE artificial intelligence toolset

The Deep Learning Cookbook helps with answering such questions as "Which deep learning framework do I need?" and "Which GPU should I choose, and how many?" But there is a lot more that HPE makes available, too:

- **Deep Learning Benchmarking Suite**
- **Deep Learning Performance Guide**
- **Reference Designs (**`http://hpe.to/6009DdXRF`**)**

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Training Your Teams Using the Deep Learning Institute

Keeping yourself and your team current with the fast-moving nature of deep learning is a critical component of your project. You need to start with the basics and become enmeshed in the communities that have been developed to bring this technology to where it is today.

HPE and NVIDIA are providing training for developers, data scientists, and researchers. The NVIDIA Deep Learning Institute (DLI) gives customers access to online labs and hands-on workshops that will prepare their workforces to utilize the latest deep learning techniques and software frameworks.

Understanding Sample Artificial Intelligence Solutions

You must have figured out that you don't need to do deep learning on your own. If you have a well-defined use case, it's entirely possible that HPE has already developed a sample model and solution that you can get and start to use immediately. No AI training time needed! Via reference architectures, you can get started with your AI initiatives and avoid much of the trial and error that is often associated with new endeavors.

Identifying fraudulent patterns

One of the most important deep learning services on the market revolves around protecting people from fraud. Every day, thieves steal people's credit cards and their very identities in order to get money. Financial fraud is a quickly growing problem and one that most people will probably face at some point in their lives. Think of it like this: How often does your bank contact you to tell you that they've detected possible fraudulent activity on your credit card?

For example, American Express recently notified me that one of my cards had been used at a gas station in California. I live in Missouri and wasn't traveling that week. Through a deep understanding of my normal activity, American Express was able to flag

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

a number of potentially fraudulent transactions and block some of them.

Many companies involved in finance and credit use such systems, and you may want to use something similar. The HPE product portfolio includes a fraud detection platform system, which can identify fraudulent activity via queries in seconds, reduce false positive rates, avoid costly losses, and generate a very quick return on investment.

Quickly exploring anomalies

The previous example is just one among many that can be cited as using deep learning to identify anomalous behavior. This kind of out-of-pattern detection becomes possible as data sets become larger and more populated and as deep learning algorithms are made available, learning what normal behaviors look like.

Once trained, these AI systems can then place a flag on an account or on a patient's health records so that a human can take a closer look and, if necessary, follow up with the cardholder or the patient.

Enhancing employee and business productivity with deep learning

Many people are concerned about the potential for AI and automation to replace people. Although this may be a valid concern, it increases productivity and human resources, which can be deployed elsewhere. As AI and deep learning systems become more powerful, they're able to replicate the results for certain human tasks, and they can do so far more quickly and cost effectively.

In fact, we've already seen this kind of shift with simple AI systems. When was the last time that you called a big company and got a phone tree? Companies used to have operators who answered every call. Today, decision systems have largely replaced these people. The remaining employees are then free to focus on more critical business activities and can leave the minutiae to the machines.

REMEMBER

Data is king. By using the *right* data at the *right* time, organizations can dramatically improve operations, making operational improvements that have measurable, impactful results. With the ability to analyze billions of data points in real time, deep learning solutions can have an immediate benefit on how your business does business.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

For example, can a chat-bot do a better job at handling initial or simple customer queries? Can an AI-backed robot do a more accurate job when it comes to fulfilling customer orders? Can a deep learning–centric cyber defense system improve the overall availability of your network infrastructure?

All of these examples and more are possible today. More importantly, these advances can help you delight your customers and increase your profits.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

IN THIS CHAPTER

- Deploying the best possible deep learning system
- Streamlining your data-centric journey
- Integrating AI into your environment
- Looking at some use cases

Chapter 5
Delivering Attractive Returns and Results

Getting your deep learning projects underway is easier when you have the right tools at your fingertips. In this chapter, I let you know about even more tools that HPE offers to help you successfully deploy AI in your business.

Improving Deployment Effectiveness

Building your deep learning environment requires significant knowledge and experience of the needs of this emerging field. Figuring out all of the pieces you need to make it work is challenging.

HPE has complete reference architectures that you can either use directly or use as a baseline for your own efforts, and build a deep learning environment that will solve your company's challenges.

TIP

HPE's deployment guides help you hit the ground running with constructs that enable a new research-driven architecture to take hold in your organization. With this architecture, you can start to use deep learning to, for example, identify suspects in crimes,

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

perform tissue classification in digital pathology, ID vehicle license plates, and much more.

HPE offers flexible deployment options for you to get started right away. Choose from an on-premises but cloud-like deployment model where you pay as you go, or buy the infrastructure outright — whichever option makes the most sense for your business.

Accelerating Analytics to Improve Business Operations

The marriage of high-performance computing (HPC) and AI has the potential to create offspring that are greater than the sum of their parts. As these technologies continue to converge, organizations are seeking comprehensive solutions to make AI easier to deploy and support.

One area that is of great interest is analytics. With the ability to delve more deeply into data, organizations can realize massive benefits. AI is helping financial services companies, but much more is possible.

There are myriad opportunities for analytics to improve business operations. Let's look at an industry in which the data patterns change, literally, every second: the weather. At its core, predicting the weather is a big data and analytics problem. By bringing more horsepower to bear, intelligence systems can more accurately predict weather patterns, which can save both lives and money. The earlier and more accurate weather predictions are, the easier it is for you to minimize its impact on your life.

And that's just for the general population. Now, imagine how that same improvement in prediction impacts air travel and delivery companies. With an improvement in weather prediction, these organizations can more effectively and with less impact alter their schedules and accommodate passengers.

Bringing HPC solutions with graphics processing unit (GPU) underpinnings to the AI party makes it possible for improved analytics to have a serious impact on business.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Integrating Artificial Intelligence and Deep Learning in Your Environment

Where do you start? You start with the business problem. If you don't know what problems you're trying to solve, you certainly won't be able to solve them! ***Remember:*** Your deep learning outcomes can be simple, or they can be complex.

Discovering powerful research capabilities

With your business problems defined, it's time to start discovering how you can harness AI to solve them. Here, you can start to "play." You don't necessarily need to build out a complete environment. There are plenty of very simple deep learning applications you can start to work with using just your desktop computer. Here's an excellent resource that can point you in the right direction so you can learn some of the basics behind discovering some real-world potential: `www.machinelearningmastery.com/inspirational-applications-deep-learning`.

TIP

When my kids were little, the motto was "learning through play." The same goes for everything you want to learn, AI and deep learning included. So, start playing!

Deploying a stable artificial intelligence environment

Ultimately, of course, after you've finished experimenting, you'll need to consider a more production-friendly environment than your laptop. That said, you don't need a lot to get started deploying a stable AI environment. HPE has a complete line of deep learning–friendly hardware resources, from the HPE ProLiant DL380 Gen10 Server up through the HPE SGI 8600.

Streamlining deployment

No matter which HPE platform you choose, work with HPE expert resources to make sure it's suitable for your needs. HPE's advisory services will help you choose the right platform and software to make even the most complex models work.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Seeing How Artificial Intelligence and Deep Learning Impact Industries

If you still aren't convinced that AI and deep learning are impacting entire industries, read on. But don't be discouraged if your needs are smaller in scope in comparison. Everyone has to start somewhere!

Improving healthcare outcomes

TIP

The benefits of AI in the healthcare industry are huge. Nurses will gain up to 50 percent more productivity using AI, patients will realize up to 9 percent cost savings through custom treatments, and most significant, the industry will see a $10 trillion savings through advances in drug discovery (that's *trillion* with a *t*). These predictions are based on a McKinsey study for North America healthcare cognitive computing market by technology that spans 2012 to 2022.

The overall potential outcomes with AI-guided healthcare are many and are significant:

- Increased accuracy of diagnoses
- Improved preventive medicine and public health
- Enhanced personalized care
- Accurate prediction of insurance costs
- Streamlined research and development with prediction models
- Guided drug development to deliver medications that meet public need
- Better patient outcomes overall

Accelerating cancer research

Although "cancer" as a monolithic disorder is unlikely to be defeated in the near future, various types of cancer are being battled every single day inside the silicon wafers that power AI clusters. Through powerful machine-based algorithms, researchers are constantly discovering new vectors by which to fight various

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

cancers. On the patient front, AI-based systems are delving deep into human bodies to help detect anomalies earlier in an effort to ensure that treatment can begin before the disease has progressed to an incurable stage.

Classifying tissue for cancer detection and disease prevention

Earlier in this book, I talked to you about simple machine learning–based AIs that used robust data sets to, for example, determine whether a photo that was being viewed contained a picture of a bird. This simple yes-or-no algorithm compares what it's looking at with its vast repository of learned images and makes a guess based on its prior knowledge.

Now, the same kind of power could be brought to bear in reviewing human tissue and classifying it as either "healthy" or "not healthy."

Gaining insight in manufacturing

What does AI mean for manufacturing? In short, as today's manufacturers continue to be pressed to reduce costs, they'll make further investments in automation. At the same time, they need to ensure that their lines keep running. A line that isn't running isn't making money.

Manufacturers can turn to an AI-assisted tool that allows them far greater insight into their operations than ever before. These systems are being brought to bear on complex problems, such as creating predictive maintenance schedules. With predictive maintenance, manufacturing organizations can avoid costly unplanned downtime thanks to an AI that lets management know when something is about to go wrong.

Staying safe online

Between the bird photos and political fights on the Internet lie some disturbing facts. One-third of all Internet traffic is considered malicious traffic intended to either defraud, disable, or damage a victim. New malware is being released every few minutes. And ransomware attacks continue to be massive moneymakers for criminals. That's not good.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

TIP

The speed at which attacks take place and new malware is released makes it impossible for human-driven cyber defenses alone to take action. By the time a human has discovered a breach, attackers may have made off with millions of sensitive data elements. If not today, then in the not-too-distant future, AI systems will be present in most organizations, watching potentially *billions* of data points in an effort to get in front of malicious activity.

Speed is critical in cyber defense, but not at the cost of accuracy. False positives from automated alerting systems can erode faith in systems and lead human security pros down the wrong path. Cyber defense systems will require very high levels of training to ensure that the human side of the equation doesn't end up breaking down.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

IN THIS CHAPTER

- Discovering the hardware that makes the magic happen
- Finding out why GPUs are the powerhouses of the AI world
- Generating returns for buyers with real-world AI-driven products

Chapter 6
Enhancing Your Competitive Advantage

In the previous chapters, I talk about what artificial intelligence (AI) and deep learning can do for you and provide generalities around how HPE can help. In this chapter, I jump into some specific products from HPE that will get you from zero to AI goodness in a nanosecond.

HPE Hardware Platforms

You know that you need hardware in order to operate your deep learning environment. HPE provides a broad portfolio that takes you from getting started through advancing your mature platforms to allow them to further propel your business.

HPE Edgeline EL4000

The HPE Edgeline EL4000 (see Figure 6-1) is designed to perform inference at the edge, regardless of environmental conditions. This purpose-built solution can acquire data, analyze it on an industry-standard x86 deep compute platform, drive high-performance displays for operators, and initiate control actions in real time — at the edge, removing transfer issues to result in faster insights and business agility.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

FIGURE 6-1: The HPE Edgeline EL4000.

HPE ProLiant DL380 Gen10

HPE's all-time best-selling server is the ProLiant DL380, currently in its tenth generation (see Figure 6-2). Although it's usable primarily for small-scale AI deployments, the DL380 can support NVIDIA GPUs for AI training, and you can add multiple servers to a cluster as you need to expand your machine learning capabilities.

FIGURE 6-2: The HPE ProLiant DL380 Gen10.

HPE Apollo sx40

The Apollo sx40 provides an extremely dense and powerful 1U GPU server for deep learning. The sx40 can support up to four NVIDIA GPUs, including the latest V100 GPUs with NVIDIA NVLink, allowing it to provide vastly accelerated computing in a very dense standard form factor, air-cooled package. The sx40 is ideal as an entry-level GPU server to investigate, learn, and develop deep learning skills and experience.

HPE Apollo 2000 Gen10

You know that scale is important in deep learning, but sometimes, you're okay starting a bit smaller. The HPE Apollo 2000 Gen10 system (see Figure 6-3) is a perfect fit for those times. The system supports up to four nodes per 2U chassis.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

FIGURE 6-3: The HPE Apollo 2000 Gen10 Chassis.

Offering a flexible configuration, the HPE Apollo 2000 Gen10 system supports a variety of workloads, from remote site systems to large AI-centric HPC clusters to everything in between. You can deploy the HPE Apollo 2000 Gen10 cost-effectively, starting small with a single 2U shared infrastructure and scaling out up to 80 HPE ProLiant Gen10 servers in a 42U rack.

The HPE Apollo 2000 is quite versatile and can be used for training models in the data center and serve as an inference engine on the edge. The ProLiant XL190r Gen10 Server node that goes into the Apollo 2000 chassis is a 2U half-width, dual-processor server with additional PCIe slots in multiple configurations providing support for additional expansion cards and support for two accelerators per server.

HPE Apollo 6500 Gen10

The HPE Apollo 6500 Gen10 (shown in Figure 6-4) is a reliable enterprise deep learning and HPC platform with industry-leading accelerators delivering exceptional performance for faster intelligence. With eight high-performance graphical processing units (GPUs) per server, the system provides superior performance per dollar for GPU-intensive workloads — delivering up to 125 TFLOPs single-precision compute. Purpose-built for accelerated computing, this platform features both the traditional PCIe and the new NVIDIA NVLink GPU interconnects, providing the flexibility to suit a wide variety of deep learning and HPC requirements. The NVIDIA NVLink GPU interconnect is particularly useful for deep learning workloads, characterized by heavy GPU-to-GPU communications.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

FIGURE 6-4: The HPE Apollo 6500 Gen10.

HPE SGI 8600

HPC and AI go hand in hand with one another. The HPE SGI 8600 is a clustered, high-performance computing (HPC) system intended to provide your AI endeavor with the horsepower necessary to meet even the most demanding needs.

For petaflop-scale deep learning, HPE introduced the HPE SGI 8600 (see Figure 6-5). This liquid-cooled platform delivers leading performance, density, and efficiency and is used for extremely large training exercises. There are three compute trays, with the HPE XA780i Gen10 Server as the ideal deep learning platform used to monitor and observe data processing for pattern recognition and anomaly detection. This tray features a two–central processing unit (CPU) socket node with the Intel Xeon Scalable processors with support for up to four NVIDIA Tesla graphics processing units (GPUs) with NVLink.

FIGURE 6-5: The HPE SGI 8600.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

A *petaflop* is a unit of system performance that equates to one quadrillion operations per second. AI systems need that kind of performance for real-time applications.

HPE AI storage solutions

Data is at the heart of deep learning. HPE provides different types of data storage solutions to meet the capacity and performance requirements of the various AI workloads, from pure capacity storage to pre-processing applications to training models in a GPU cluster. A hyperconverged approach offers fastest throughput with direct access to GPUs, while detached storage allows for flexibility to scale GPU and storage independently. HPE provides flexibility and choice of storage options to fit your infrastructure requirements.

WekaIO Matrix

WekaIO Matrix is a flash-optimized parallel file system for large training sets and large GPU clusters. The WekaIO solution provides high performance and low latency for faster data set processing. With support for multiple protocols, such as POSIX, S3, and HDFS, automatic data tiering, and cloud bursting capabilities, WekaIO Matrix can flexibly support the entire deep learning dataflow and scale for performance and capacity growth as system requirements evolve.

DDN

A hybrid scale-out parallel file system solution, DDN high-performance storage solutions scale from test-bed to production with a single platform. The DDN storage and flash solutions can support the entire deep learning dataflow, including providing capacity, ingest capability, and flash burst buffer technology for performance and responsiveness. By providing an integrated hybrid environment and using flash and hard drives, DDN provides the ability to define a cost-effective scaling strategy based on current requirements for either capacity or performance.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

HPE Apollo 4200 with Qumulo QF2

The HPE Apollo 4200 system with Qumulo QF2 offers an end-to-end scale-out NAS solution for unstructured data, offering capacity storage as well as linear performance scalability. Data reliability is provided with block-level erasure coding and built-in analytics, allowing administrators to pinpoint problems and control how storage is used. Cloud capabilities with continuous replication and snapshots offer additional scalability with cloud economics.

HPE Scalable Object Storage with Scality RING

A dense storage system for actively archiving large data sets, HPE Scalable Object Storage with Scality RING is a more economic storage tier as data gets colder.

HPE Scalable Object Storage with Scality RING provides the ideal large-scale active archive storage tier for deep learning data. Deep learning often requires large amounts of data, which then must be labeled, processed, and iterated upon, which further increases the amount of storage that is required post-processing. HPE Scalable Object Storage is an economic, dense storage solution that provides reasonable access speed while serving as the base tier archive for all deep learning applications.

HPE Data Management Framework

Deep learning data is not static. It gets cleaned, labeled, processed, transformed, and iterated upon. This makes data management essential, because data files must be not only stored, but also collected and moved to serve compute requirements. The HPE Data Management Framework, working in concert with AI storage platforms, simplifies data workflows, using an automated policy engine to move data to the right place at the right time for processing requirements, and then moves the results of processing to the appropriate storage infrastructure to reduce overall cycle times.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

How Accelerated Graphics Processing Unit Compute Impacts Learning Speed

Why use expensive GPUs when you could just use a bunch of less expensive CPUs to operate your deep learning environment? It comes down to how each is architected and operates.

There are a number of key architectural differences between AI and typical enterprise applications. Enterprise applications work well with standard CPUs because those applications aren't generally all that parallel in nature. AI, on the other hand, can be a massively parallel operation. As a result, you could easily flood an entire rack of CPU-based servers with an AI problem and still not get the results you'd see from a good GPU.

Why? It's all about the cores. CPUs might have, say, 20 cores. GPUs, on the other hand, can have thousands. And because each individual workload can be pretty small, you don't need massive cores to complete individual tasks. GPUs are purpose-built to handle complex floating-point mathematics equations with great precision.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Chapter 7
Ten (Or So) Ways HPE Can Help You Get Started

It's clear that getting started with artificial intelligence, deep learning, and machine learning should be on your short list of learning goals. Here are ten (or so) ways that HPE can help you get underway:

- **Turn to HPE for help in making deep learning work for you.** HPE has several tools to help you overcome your IT challenges and leverage the latest deep learning technologies to gain competitive advantage.

  With the right mix of hardware, software, services, and training, any organization can harness the power of AI. Check out the following for more information:

  - **HPE Demystifies Deep Learning For Faster Intelligence:** `www.nextplatform.com/2017/10/25/hpe-demystifies-deep-learning-faster-intelligence/`
  - **HPE Apollo 6500 for Deep Learning:** `www.youtube.com/watch?v=wkI4CFKWxFs`

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

- **Transformation workshop:** https://h20195.www2.hpe.com/v2/Getdocument.aspx?docname=a00042611enw
- **Transformation services:** www.hpe.com/us/en/services/consulting.html#servicesPortfolio
- **HPE Pointnext services:** www.hpe.com/services

» **See how HPE can help you figure out your optimal recipe for configuring your hardware and software.** You can use tools like the Deep Learning Cookbook, which includes a benchmarking suite, the performance analysis tool, and reference designs. You can also further educate yourself by checking out the Deep Learning Workshops. Go to http://dlpg.labs.hpe.com.

» **Discover HPE's sample AI solutions from a variety of verticals, which can help ease you into the world of deep learning and assist you in getting results from your investment into the technology.**

» **Learn from the comprehensive resources that HPE has made available for you, all intended to help you understand how deep learning can be used in your organization:**

- **Deep Learning Workshops:** https://h10076.www1.hpe.com/ww/en/training/solution/deep_learning_AI.html
- **Deep Learning Cookbook:** https://developer.hpe.com/platform/hpe-deep-learning-cookbook/home

» **Find out how HPE is establishing itself as a leader in deep learning technology and giving back to the community in an effort to accelerate deployment and standardization.**

- **DL benchmarking suite:** https://community.hpe.com/t5/Behind-the-scenes-Labs/The-Deep-Learning-Benchmarking-Suite-now-available/ba-p/6989905#.WoIF2SXwaUk
- **Comparative analysis of different DL solutions:** http://on-demand.gputechconf.com/gtc-eu/2017/presentation/23245-sorin-cheran-hpes-deep-learning-cookbook.pdf

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

- **Let HPE help you with your research via two key resources:**
  - **HPE AI Innovation Center:** Schedule time at the HPE AI Innovation Center, which is designed for longer-term research projects. The Innovation Center serves as a platform for research collaboration between universities, enterprises on the cutting edge of AI research, and HPE researchers. The centers — located in Houston, Texas; Palo Alto, California; and Grenoble, France — give researchers for academia and enterprises access to infrastructure and tools to continue research initiatives
  - **Enhanced HPE Centers of Excellence (CoEs):** Visit an Enhanced HPE CoE for help from the best minds in AI. CoEs are designed to assist IT departments and data scientists who are looking to accelerate their deep learning applications and realize better return on investment (ROI) from their deep learning deployments in the near term. The HPE CoEs offers select customers access to the latest technology and expertise, including the latest NVIDIA GPUs on HPE systems. The current CoEs are spread across five locations: Houston, Texas; Palo Alto, California; Tokyo, Japan; Bangalore, India; and Grenoble, France.
- **See how HPE can help your AI project on the financial front.** Can't quite afford the upfront spend you might associate with a deep learning project? Discover HPE GreenLake, HPE's on-premises pay-as-you-go consumption model. Pay only for what you use. Whether it's infrastructure, real-time analytics, big data, backup, or protecting apps and data — business leaders now widely prefer to consume the outcomes they want. Visit HPE GreenLake Flex Capacity at `www.hpe.com/services/flexiblecapacity`.
- **Discover more help from HPE in the form of the HPE Data Management Framework.** The HPE Data Management Framework, working in concert with AI storage platforms, simplifies data workflows, using an automated policy engine to move data to the right place at the right time for processing requirements, and then moves the results of processing to the appropriate storage infrastructure to reduce overall cycle times.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

- **Find out how HPE is going beyond its own walls to help customers deploy the most comprehensive solution.** Discover HPE's broad partner ecosystem for deep learning. From fraud detection solutions to custom solutions built by HPE Pointnext and HPE partners, there's a solution for all of your deep learning needs.

You're now ready for your own deep learning journey to success. You can

- Contact an HPE sales rep to learn how you can start on your journey today.
- Download articles to augment your existing knowledge.
- Visit the HPE Deep Learning website (`www.hpe.com/info/deep-learning`).

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

About Hewlett Packard Enterprise

Hewlett Packard Enterprise is an industry-leading technology company that enables customers to go further, faster. With the industry's most comprehensive portfolio, spanning the cloud to the data center to workplace applications, our technology and services help customers around the world make IT more efficient, more productive and more secure.

These materials are © 2018 John Wiley & Sons, Inc. Any dissemination, distribution, or unauthorized use is strictly prohibited.

Adopt deep learning for faster business outcomes

Today, there are numerous examples of deep learning applications in our personal lives. However, adopting deep learning in a business context has its inherent challenges. This book helps you understand the fundamentals of deep learning, how to overcome adoption challenges, and how HPE can help deliver deep learning solutions for faster and better business outcomes from your data.

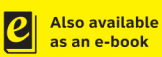
Inside...

- Identify challenges hindering deep learning adoption
- See how HPE demystifies deep learning principles
- Build deep learning models faster
- Scale and integrate models efficiently
- Find examples of successful deep learning implementations
- Get tailored products and solutions from HPE

Scott D. Lowe is CEO of ActualTech Media. Scott spent ten years as a CIO for various organizations. With a breadth of experience, Scott brings insight to both technical and decision-making audiences. Scott is still a practicing consultant, providing strategic IT direction to his clients.

Cover Image: © HPE

Go to Dummies.com®
for videos, step-by-step photos, how-to articles, or to shop!



ISBN: 978-1-119-48458-5
Not for resale

WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley’s ebook EULA.