Distributed Deep Learning Framework Over Spark

Take a systematic approach to understanding the fundamentals of machine learning and deep learning from the ground up and how they are applied in practice. You will use this comprehensive guide for building and deploying learning models to address complex use cases while leveraging the computational resources of Google Cloud Platform. Author Ekaba Bisong shows you how machine learning tools and techniques are used to predict or classify events based on a set of interactions between variables known as features or attributes in a particular dataset. He teaches you how deep learning extends the machine learning algorithm of neural networks to learn complex tasks that are difficult for computers to perform, such as recognizing faces and understanding languages. And you will know how to leverage cloud computing to accelerate data science and machine learning deployments. Building Machine Learning and Deep Learning Models on Google Cloud Platform is divided into eight parts that cover the fundamentals of machine learning and deep learning, the concept of data science and cloud services, programming for data science using the Python stack, Google Cloud Platform (GCP) infrastructure and products, advanced analytics on GCP, and deploying end-to-end machine learning solution pipelines on GCP. What You'll Learn Understand the principles and fundamentals of machine learning and deep learning, the algorithms, how to use them, when to use them, and how to interpret your results Know the programming concepts relevant to machine and deep learning design and development using the Python stack Build and interpret machine and deep learning models Use Google Cloud Platform tools and services to develop and deploy large-scale machine learning and deep learning products Be aware of the different facets and design choices to consider when modeling a learning problem Productionalize machine learning models into software products Who This Book Is For Beginners to the practice of data science and applied machine learning, data scientists at all levels, machine learning engineers, Google Cloud Platform data engineers/architects, and software developers

As an important enabler for changing people's lives, advances in artificial intelligence (AI)-based applications and services are on the rise, despite being hindered by efficiency and latency issues. By focusing on deep learning as the most representative technique of AI, this book provides a comprehensive overview of how AI services are being applied to the network edge near the data sources, and demonstrates how AI and edge computing can be mutually beneficial. To do so, it introduces and discusses: 1) edge intelligence and intelligent edge; and 2) their implementation methods and enabling technologies, namely AI training and inference in the customized edge computing framework. Gathering essential information previously scattered across the communication, networking, and AI areas, the book can help readers to understand the connections between key enabling technologies, e.g. a) AI applications in edge; b) AI inference in edge; c) AI training for edge; d) edge computing for AI; and e) using AI to optimize edge. After identifying these five aspects, which are essential for the fusion of edge computing and AI, it discusses current challenges and outlines future trends in achieving more pervasive and fine-grained intelligence with the aid of edge computing.

High-Performance Computing for Big Data: Methodologies and Applications explores emerging high-performance architectures for data-intensive applications, novel efficient analytical strategies to boost data processing, and cutting-edge applications in diverse fields, such as machine learning, life science, neural networks, and neuromorphic engineering. The book is organized into two main sections. The first section covers Big Data architectures, including cloud computing systems, and heterogeneous accelerators. It also covers emerging 3D IC design principles for memory architectures and devices. The second section of the book illustrates emerging and practical applications of Big Data across several domains, including bioinformatics, deep learning, and neuromorphic engineering. Features Covers a wide range of Big Data architectures, including distributed systems like Hadoop/Spark Includes accelerator-based approaches for big data applications such as GPU-based acceleration techniques, and hardware acceleration such as FPGA/CGRA/ASICs Presents emerging memory architectures and devices such as NVM, STT- RAM, 3D IC design principles Describes advanced algorithms for different big data application domains Illustrates novel analytics techniques for Big Data applications, scheduling, mapping, and partitioning methodologies Featuring contributions from leading experts, this book presents state-of-the-art research on the methodologies and applications of high-performance computing for big data applications. About the Editor Dr. Chao Wang is an Associate Professor in the School of Computer Science at the University of Science and Technology of China. He is the Associate Editor of ACM Transactions on Design Automations for Electronics Systems (TODAES), Applied Soft Computing, Microprocessors and Microsystems, IET Computers & Digital Techniques, and International Journal of Electronics. Dr. Chao Wang was the recipient of Youth Innovation Promotion Association, CAS, ACM China Rising Star Honorable Mention (2016), and best IP nomination of DATE 2015. He is now on the CCF Technical Committee on Computer Architecture, CCF Task Force on Formal Methods. He is a Senior Member of IEEE, Senior Member of CCF, and a Senior Member of ACM.

This illuminating text/reference surveys the state of the art in data science, and provides practical guidance on big data analytics. Expert perspectives are provided by authoritative researchers and practitioners from around the world, discussing research developments and emerging trends, presenting case studies on helpful frameworks and innovative methodologies, and suggesting best practices for efficient and effective data analytics. Features: reviews a framework for fast data applications, a technique for complex event processing, and agglomerative approaches for the partitioning of networks; introduces a unified approach to data modeling and management, and a distributed computing perspective on interfacing physical and cyber worlds; presents techniques for machine learning for big data, and identifying duplicate records in data repositories; examines enabling technologies and tools for data mining; proposes frameworks for data extraction, and adaptive decision making and social media analysis.

A solution-based guide to put your deep learning models into production with the power of Apache Spark Key Features Discover practical recipes for distributed deep learning with Apache Spark Learn to use libraries such as Keras and TensorFlow Solve problems in order to train your deep learning models on Apache Spark Book Description With deep learning gaining rapid mainstream adoption in modern-day industries, organizations are looking for ways to unite popular big data tools with highly efficient deep learning libraries. As a result, this will help deep learning models train with higher efficiency and speed. With the help of the Apache Spark Deep Learning Cookbook, you'll work through specific recipes to generate outcomes for deep learning algorithms, without getting bogged down in theory. From setting up Apache Spark for deep learning to implementing types of neural net, this book tackles both common and not so common problems to perform deep learning on a distributed environment. In addition to this, you'll get access to deep learning code within Spark that can be reused to answer similar problems or tweaked to answer slightly different problems. You will also learn how to stream and cluster your data with Spark. Once you have got to grips with the basics, you'll explore how to implement and deploy deep learning models, such as Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) in Spark, using popular libraries such as TensorFlow and Keras. By the end of the book, you'll have the expertise to train and deploy efficient deep learning models on Apache Spark. What you will learn Set up a fully functional Spark environment Understand practical machine learning and deep learning concepts Apply built-in machine learning libraries within Spark Explore libraries that are compatible with TensorFlow and Keras Explore NLP models such as Word2vec and TF-IDF on Spark Organize dataframes for deep learning evaluation Apply testing and training modeling to ensure accuracy Access readily available code that may be reusable Who this book is for If you're looking for a practical and highly useful resource for implementing efficiently distributed deep learning models with Apache Spark, then the Apache Spark Deep Learning Cookbook is for you. Knowledge of the core machine learning concepts and a basic understanding of the Apache Spark framework is required to get the best out of this book. Additionally, some programming knowledge in Python is a plus.

While Computer Security is a broader term which incorporates technologies, protocols, standards and policies to ensure the security of the computing systems including the computer hardware, software and the information stored in it, Cyber Security is a specific, growing field to protect computer networks (offline and online) from unauthorized access, botnets, phishing scams, etc. Machine learning is a branch of Computer Science which enables computing machines to adopt new behaviors on the basis of observable and verifiable data and information. It can be applied to ensure the security of the computers and the information by detecting anomalies using data mining and other such techniques. This book will be an invaluable resource to understand the importance of machine learning and data mining in establishing computer and cyber security. It emphasizes important security aspects associated with computer and cyber security along with the analysis of machine learning and data mining based solutions. The book also highlights the future research domains in which these solutions can be applied. Furthermore, it caters to the needs of IT professionals, researchers, faculty members, scientists, graduate students, research scholars and software developers who seek to carry out research and develop combating solutions in the area of cyber security using machine learning based approaches. It is an extensive source of information for the readers belonging to the field of Computer Science and Engineering, and Cyber Security professionals. Key Features: This book contains examples and illustrations to demonstrate the principles, algorithms, challenges and applications of machine learning and data mining for computer and cyber security. It showcases important security aspects and current trends in the field. It provides an insight of the future research directions in the field. Contents of this book help to prepare the students for exercising better defense in terms of understanding the motivation of the attackers and how to deal with and mitigate the situation using machine learning based approaches in better manner.

This book presents recent machine learning paradigms and advances in learning analytics, an emerging research discipline concerned with the collection, advanced processing, and extraction of useful information from both educators' and learners' data with the goal of improving education and learning systems. In this context, internationally respected researchers present various aspects of learning analytics and selected application areas, including: • Using learning analytics to measure student engagement, to quantify the learning experience and to facilitate self-regulation; • Using learning analytics to predict student performance; • Using learning analytics to create learning materials and educational courses; and • Using learning analytics as a tool to support learners and educators in synchronous and asynchronous eLearning. The book offers a valuable asset for professors, researchers, scientists, engineers and students of all disciplines. Extensive bibliographies at the end of each chapter guide readers to probe further into their application areas of interest.

Whether you're a software engineer aspiring to enter the world of deep learning, a veteran data scientist, or a hobbyist with a simple dream of making the next viral AI app, you might have wondered where to begin. This step-by-step guide teaches you how to build practical deep learning applications for the cloud, mobile, browsers, and edge devices using a hands-on approach. Relying on years of industry experience transforming deep learning research into award-winning applications, Anirudh Koul, Siddha Ganju, and Meher Kasam guide you through the process of converting an idea into something that people in the real world can use. Train, tune, and deploy computer vision models with Keras, TensorFlow, Core ML, and TensorFlow Lite Develop AI for a range of devices including Raspberry Pi, Jetson Nano, and Google Coral Explore fun projects, from Silicon Valley's Not Hotdog app to 40+ industry case studies Simulate an autonomous car in a video game environment and build a miniature version with reinforcement learning Use transfer learning to train models in minutes Discover 50+ practical tips for maximizing model accuracy and speed, debugging, and scaling to millions of users

"We finally have the definitive treatise on PyTorch! It covers the basics and abstractions in great detail. I hope this book becomes your extended reference document." —Soumith Chintala, co-creator of PyTorch Key Features Written by PyTorch's creator and key contributors Develop deep learning models in a familiar Pythonic way Use PyTorch to build an image classifier for cancer detection Diagnose problems with your neural network and improve training with data augmentation Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications. About The Book Every other day we hear about new ways to put deep learning to good use: improved medical imaging, accurate credit card fraud detection, long range weather forecasting, and more. PyTorch puts these superpowers in your hands. Instantly familiar to anyone who knows Python data tools like NumPy and Scikit-learn, PyTorch simplifies deep learning without sacrificing advanced features. It's great for building quick models, and it scales smoothly from laptop to enterprise. Deep Learning with PyTorch teaches you to create deep learning and neural network systems with PyTorch. This practical book gets you to work right away building a tumor image classifier from scratch. After covering the basics, you'll learn best practices for the entire deep learning pipeline, tackling advanced projects as your PyTorch skills become more sophisticated. All code samples are easy to explore in downloadable Jupyter notebooks. What You Will Learn Understanding deep learning data structures such as tensors and neural networks Best practices for the PyTorch Tensor API, loading data in Python, and visualizing results Implementing modules and loss functions Utilizing pretrained models from PyTorch Hub Methods for training networks with limited inputs Sifting through unreliable results to diagnose and fix problems in your neural network Improve your results with augmented data, better model architecture, and fine tuning This Book Is Written For For Python programmers with an interest in machine learning. No experience with PyTorch or other deep learning frameworks is required. About The Authors Eli Stevens has worked in Silicon Valley for the past 15 years as a software engineer, and the past 7 years as Chief Technical Officer of a startup making medical device software. Luca Antiga is co-founder and CEO of an AI engineering company located in Bergamo, Italy, and a regular contributor to PyTorch. Thomas Viehmann is a Machine Learning and PyTorch speciality trainer and consultant based in Munich, Germany and a PyTorch core developer. Table of Contents PART 1 - CORE PYTORCH 1 Introducing deep learning and the PyTorch Library 2 Pretrained networks 3 It starts with a tensor 4 Real-world data representation using tensors 5 The mechanics of learning 6 Using a neural network to fit the data 7 Telling birds from airplanes: Learning from images 8 Using convolutions to generalize PART 2 - LEARNING FROM IMAGES IN THE REAL WORLD: EARLY DETECTION OF LUNG CANCER 9 Using PyTorch to fight cancer 10 Combining data sources into a unified dataset 11 Training a classification model to detect suspected tumors 12 Improving training with metrics and augmentation 13 Using segmentation to find suspected nodules 14 End-to-end nodule analysis, and where to go next PART 3 - DEPLOYMENT 15 Deploying to production

This book constitutes the refereed post-conference proceedings of the 5th Russian Supercomputing Days, RuSCDays 2019, held in Moscow, Russia, in September 2019. The 60 revised full papers presented were carefully reviewed and selected from 127 submissions. The papers are organized in the following topical sections: parallel algorithms; supercomputer simulation; HPC, BigData, AI: architectures, technologies, tools; and distributed and cloud computing.

This book presents the proceedings of the 24th European Conference on Artificial Intelligence (ECAI 2020), held in Santiago de Compostela, Spain, from 29 August to 8 September 2020. The conference was postponed from June, and much of it conducted online due to the COVID-19 restrictions. The conference is one of the principal occasions for researchers and practitioners of AI to meet and discuss the latest trends and challenges in all fields of AI and to demonstrate innovative applications and uses of advanced AI technology. The book also includes the proceedings of the 10th Conference on Prestigious Applications of Artificial Intelligence (PAIS 2020) held at the same time. A record number of more than 1,700 submissions was received for ECAI 2020, of which 1,443 were reviewed. Of these, 361 full-papers and 36 highlight papers were accepted (an acceptance rate of 25% for full-papers and 45% for highlight papers). The book is divided into three sections: ECAI full papers; ECAI highlight papers; and PAIS papers. The topics of these papers cover all aspects of AI, including Agent-based and Multi-agent Systems; Computational Intelligence; Constraints and Satisfiability; Games and Virtual Environments; Heuristic Search; Human Aspects in AI; Information Retrieval and Filtering; Knowledge Representation and Reasoning; Machine Learning; Multidisciplinary Topics and Applications; Natural Language Processing; Planning and Scheduling; Robotics; Safe, Explainable, and Trustworthy AI; Semantic Technologies; Uncertainty in AI; and Vision. The book will be of interest to all those whose work involves the use of AI technology.

Learn how to use, deploy, and maintain Apache Spark with this comprehensive guide, written by the creators of the open-source cluster-computing framework. With an emphasis on improvements and new features in Spark 2.0, authors Bill Chambers and Matei Zaharia break down Spark topics into distinct sections, each with unique goals. You'll explore the basic operations and common functions of Spark's structured APIs, as well as Structured Streaming, a new high-level API for building end-to-end streaming applications. Developers and system administrators will learn the fundamentals of monitoring, tuning, and debugging Spark, and explore machine learning techniques and scenarios for employing MLlib, Spark's scalable machine-learning library. Get a gentle overview of big data and Spark Learn about DataFrames, SQL, and Datasets—Spark's core APIs—through worked examples Dive into Spark's low-level APIs, RDDs, and execution of SQL and DataFrames Understand how Spark runs on a cluster Debug, monitor, and tune Spark clusters and applications Learn the power of Structured Streaming, Spark's stream-processing engine Learn how you can apply MLlib to a variety of problems, including classification or recommendation

Deep neural networks have become popular for solving machine learning problems in the field of computer vision. Although computers have reached parity in the task of image classification in machine learning competitions, the task of mining massive training data often takes expensive hardware a long time to process. Distributed protocol for model training can be attractive because less powerful distributed nodes are cheaper to operate than specialized high-performance cluster. Stochastic gradient descent (SGD) is a popular optimizer at the heart of many deep learning systems. To investigate the performance of distributed asynchronous SGD, Tensorflow deep learning framework was tested with Downpour SGD and Delay Compensated SGD to see effect of model training in typical commercial environments. Experimental results show that both Downpour and Delay Compensated SGD are viable protocols for distributed deep learning.

The 4-volume set LNCS 11632 until LNCS 11635 constitutes the refereed proceedings of the 5th International Conference on Artificial Intelligence and Security, ICAIS 2019, which was held in New York, USA, in July 2019. The conference was formerly called "International Conference on Cloud Computing and Security" with the acronym ICCCS. The total of 230 full papers presented in this 4-volume proceedings was carefully reviewed and selected from 1529 submissions. The papers were organized in topical sections as follows: Part I: cloud computing; Part II: artificial intelligence; big data; and cloud computing and security; Part III: cloud computing and security; information hiding; IoT security; multimedia forensics; and encryption and cybersecurity; Part IV: encryption and cybersecurity.

Build, implement and scale distributed deep learning models for large-scale datasets About This Book Get to grips with the deep learning concepts and set up Hadoop to put them to use Implement and parallelize deep learning models on Hadoop's YARN framework A comprehensive tutorial to distributed deep learning with Hadoop Who This Book Is For If you are a data scientist who wants to learn how to perform deep learning on Hadoop, this is the book for you. Knowledge of the basic machine learning concepts and some understanding of Hadoop is required to make the best use of this book. What You Will Learn Explore Deep Learning and various models associated with it Understand the challenges of implementing distributed deep learning with Hadoop and how to overcome it Implement Convolutional Neural Network (CNN) with deeplearning4j Delve into the implementation of Restricted Boltzmann Machines (RBM) Understand the mathematical explanation for implementing Recurrent Neural Networks (RNN) Get hands on practice of deep learning and their implementation with Hadoop. In Detail This book will teach you how to deploy large-scale dataset in deep neural networks with Hadoop for optimal performance. Starting with understanding what deep learning is, and what the various models associated with deep neural networks are, this book will then show you how to set up the Hadoop environment for deep learning. In this book, you will also learn how to overcome the challenges that you face while implementing distributed deep learning with large-scale unstructured datasets. The book will also show you how you can implement and parallelize the widely used deep learning models such as Deep Belief Networks, Convolutional Neural Networks, Recurrent Neural Networks, Restricted Boltzmann Machines and autoencoder using the popular deep learning library deeplearning4j. Get in-depth mathematical explanations and visual representations to help you understand the design and implementations of Recurrent Neural network and Denoising AutoEncoders with deeplearning4j. To give you a more practical perspective, the book will also teach you the implementation of large-scale video processing, image processing and natural language processing on Hadoop. By the end of this book, you will know how to deploy various deep neural networks in distributed systems using Hadoop. Style and approach

This book takes a comprehensive, step-by-step approach to implement efficient deep learning models on Hadoop. It starts from the basics and builds the readers' knowledge as they strengthen their understanding of the concepts. Practical examples are included in every step of the way to supplement the theory.

This volume constitutes the proceedings of the 7th International Conference on BIGDATA 2018, held as Part of SCF 2018 in Seattle, WA, USA in June 2018. The 22 full papers together with 10 short papers published in this volume were carefully reviewed and selected from 97 submissions. They are organized in topical sections such as Data analysis, data as a service, services computing, data conversion, data storage, data centers, dataflow architectures, data compression, data exchange, data modeling, databases, and data management.

This book presents the combined proceedings of the 12th KIPS International Conference on Ubiquitous Information Technologies and Applications (CUTE 2017) and the 9th International Conference on Computer Science and its Applications (CSA2017), both held in Taichung, Taiwan, December 18 - 20, 2017. The aim of these two meetings was to promote discussion and interaction among academics, researchers and professionals in the field of ubiquitous computing technologies. These proceedings reflect the state of the art in the development of computational methods, involving theory, algorithms, numerical simulation, error and uncertainty analysis and novel applications of new processing techniques in engineering, science, and other disciplines related to ubiquitous computing. James J. (Jong Hyuk) Park received Ph.D. degrees in Graduate School of Information Security from Korea University, Korea and Graduate School of Human Sciences from Waseda University, Japan. From December, 2002 to July, 2007, Dr. Park had been a research scientist of R&D Institute, Hanwha S&C Co., Ltd., Korea. From September, 2007 to August, 2009, He had been a professor at the Department of Computer Science and Engineering, Kyungnam University, Korea. He is now a professor at the Department of Computer Science and Engineering and Department of Interdisciplinary Bio IT Materials, Seoul National University of Science and Technology (SeoulTech), Korea. Dr. Park has published about 200 research papers in international journals and conferences. He has been serving as chair, program committee, or organizing committee chair for many international conferences and workshops. He is a steering chair of international conferences – MUE, FutureTech, CSA, CUTE, UCAWSN, World IT Congress-Jeju. He is editor-in-chief of Human-centric Computing and Information Sciences (HCIS) by Springer, The Journal of Information Processing Systems (JIPS) by KIPS, and Journal of Convergence (JoC) by KIPS CSWRG. He is Associate Editor / Editor of 14 international journals including JoS, JNCA, SCN, CJ, and so on. In addition, he has been serving as a Guest Editor for international journals by some publishers: Springer, Elsevier, John Wiley, Oxford Univ. press, Emerald, Inderscience, MDPI. He got the best paper awards from ISA-08 and ITCS-11 conferences and the outstanding leadership awards from IEEE HPCC-09, ICA3PP-10, IEE ISPA-11, PDCAT-11, IEEE AINA-15. Furthermore, he got the outstanding research awards from the SeoulTech, 2014. His research interests include IoT, Human-centric Ubiquitous Computing, Information Security, Digital Forensics, Vehicular Cloud Computing, Multimedia Computing, etc. He is a member of the IEEE, IEEE Computer Society, KIPS, and KMMS. Vincenzo Loia (BS '85, MS '87, PhD '89) is Full Professor of Computer Science. His research interests include Intelligent Agents, Ambient intelligence, Computational Intelligence. Currently he is Founder & Editor-in-chief of "Ambient Intelligence and Humanized Computing", and Co-Editor-in-Chief of "Softcomputing", Springer-Verlag. He is Chair of the Task Forces "Intelligent Agents" and "Ambient Intelligence" IEEE CIS ETTC. He has been Chair the Emergent Technical Committe "Emergent Technology", IEEE CIS Society and Vice-Chair of Intelligent Systems Applications Technical Committee. He has been author of more than 200 scientific works, Editor/co-editor of 4 Books, 64 journal papers, 25 book chapters, and 100 conference papers. He is Senior member of the IEEE, Associate Editor of IEEE Transactions on Industrial Informatics, and Associate Editor of IEEE Transactions on Systems, Man, and Cybernetics: Systems. Many times reviewers for national and international projects, Dr. Loia is active in the research domain of agents, ambient intelligence, computational intelligence, smartgrids, distributed platform for enrich added value. Gangman Yi in Computer Sciences at Texas A&M University, USA in 2007, and doctorate in Computer Sciences at Texas A&M University, USA in 2011. In May 2011, he joined System S/W group in Samsung Electronics, Suwon, Korea. He joined the Department of Computer Science & Engineering, Gangneung-Wonju National University, Korea, since March 2012. Dr. Yi has been researched in an interdisciplinary field of researches. His research focuses especially on the development of computational methods to improve understanding of biological systems and its big data. Dr. Yi actively serves as a managing editor and reviewer for international journals, and chair of international conferences and workshops. Yunsick Sung received his B.S. degree in division of electrical and computer engineering from Pusan National University, Busan, Korea, in 2004, his M.S. degree in computer engineering from Dongguk University, Seoul, Korea, in 2006, and his Ph.D. degree in game engineering from Dongguk University, Seoul, Korea, in 2012. He was employed as a member of the researcher at Samsung Electronics between 2006 and 2009. He was the plural professor at Shinheung College in 2009 and at Dongguk University in 2010. His main research interests are many topics in brain-computer Interface, programming by demonstration, ubiquitous computing and reinforcement learning. His Journal Service Experiences is Associate Editor at Human-centric Computing and Information Sciences, Springer (2015- Current).

Design, implement, and deliver successful streaming applications, machine learning pipelines and graph applications using Spark SQL API About This Book Learn about the design and implementation of streaming applications, machine learning pipelines, deep learning, and large-scale graph processing applications using Spark SQL APIs and Scala. Learn data exploration, data munging, and how to process structured and semi-structured data using real-world datasets and gain hands-on exposure to the issues and challenges of working with noisy and "dirty" real-world data. Understand design considerations for scalability and performance in web-scale Spark application architectures. Who This Book Is For If you are a developer, engineer, or an architect and want to learn how to use Apache Spark in a web-scale project, then this is the book for you. It is assumed that you have prior knowledge of SQL querying. A basic programming knowledge with Scala, Java, R, or Python is all you need to get started with this book. What You Will Learn Familiarize yourself with Spark SQL programming, including working with DataFrame/Dataset API and SQL Perform a series of hands-on exercises with different types of data sources, including CSV, JSON, Avro, MySQL, and MongoDB Perform data quality checks, data visualization, and basic statistical analysis tasks Perform data munging tasks on publically available datasets Learn how to use Spark SQL and Apache Kafka to build streaming applications Learn key performance-tuning tips and tricks in Spark SQL applications Learn key architectural components and patterns in large-scale Spark SQL applications In Detail In the past year, Apache Spark has been increasingly adopted for the development of distributed applications. Spark SQL APIs provide an optimized interface that helps developers build such

applications quickly and easily. However, designing web-scale production applications using Spark SQL APIs can be a complex task. Hence, understanding the design and implementation best practices before you start your project will help you avoid these problems. This book gives an insight into the engineering practices used to design and build real-world, Spark-based applications. The book's hands-on examples will give you the required confidence to work on any future projects you encounter in Spark SQL. It starts by familiarizing you with data exploration and data munging tasks using Spark SQL and Scala. Extensive code examples will help you understand the methods used to implement typical use-cases for various types of applications. You will get a walkthrough of the key concepts and terms that are common to streaming, machine learning, and graph applications. You will also learn key performance-tuning details including Cost Based Optimization (Spark 2.2) in Spark SQL applications. Finally, you will move on to learning how such systems are architected and deployed for a successful delivery of your project. Style and approach This book is a hands-on guide to designing, building, and deploying Spark SQL-centric production applications at scale.

This two-volume set (CCIS 1367-1368) constitutes reviewed and selected papers from the 10th International Advanced Computing Conference, IACC 2020, held in December 2020. The 65 full papers and 2 short papers presented in two volumes were thoroughly reviewed and selected from 286 submissions. The papers are organized in the following topical sections: Application of Artificial Intelligence and Machine Learning in Healthcare; Using Natural Language Processing for Solving Text and Language related Applications; Using Different Neural Network Architectures for Interesting applications; ?Using AI for Plant and Animal related Applications.- Applications of Blockchain and IoT.- Use of Data Science for Building Intelligence Applications; Innovations in Advanced Network Systems; Advanced Algorithms for Miscellaneous Domains; New Approaches in Software Engineering.

Deep Learning and Parallel Computing Environment for Bioengineering Systems delivers a significant forum for the technical advancement of deep learning in parallel computing environment across bio-engineering diversified domains and its applications. Pursuing an interdisciplinary approach, it focuses on methods used to identify and acquire valid, potentially useful knowledge sources. Managing the gathered knowledge and applying it to multiple domains including health care, social networks, mining, recommendation systems, image processing, pattern recognition and predictions using deep learning paradigms is the major strength of this book. This book integrates the core ideas of deep learning and its applications in bio engineering application domains, to be accessible to all scholars and academicians. The proposed techniques and concepts in this book can be extended in future to accommodate changing business organizations' needs as well as practitioners' innovative ideas. Presents novel, in-depth research contributions from a methodological/application perspective in understanding the fusion of deep machine learning paradigms and their capabilities in solving a diverse range of problems Illustrates the state-of-the-art and recent developments in the new theories and applications of deep learning approaches applied to parallel computing environment in bioengineering systems Provides concepts and technologies that are successfully used in the implementation of today's intelligent data-centric critical systems and multi-media Cloud-Big data

Deploy deep learning solutions in production with ease using TensorFlow. You'll also develop the mathematical understanding and intuition required to invent new deep learning architectures and solutions on your own. Pro Deep Learning with TensorFlow provides practical, hands-on expertise so you can learn deep learning from scratch and deploy meaningful deep learning solutions. This book will allow you to get up to speed quickly using TensorFlow and to optimize different deep learning architectures. All of the practical aspects of deep learning that are relevant in any industry are emphasized in this book. You will be able to use the prototypes demonstrated to build new deep learning applications. The code presented in the book is available in the form of iPython notebooks and scripts which allow you to try out examples and extend them in interesting ways. You will be equipped with the mathematical foundation and scientific knowledge to pursue research in this field and give back to the community. What You'll Learn Understand full stack deep learning using TensorFlow and gain a solid mathematical foundation for deep learning Deploy complex deep learning solutions in production using TensorFlow Carry out research on deep learning and perform experiments using TensorFlow Who This Book Is For Data scientists and machine learning professionals, software developers, graduate students, and open source enthusiasts

This book presents deep learning techniques, concepts, and algorithms to classify and analyze big data. Further, it offers an introductory level understanding of the new programming languages and tools used to analyze big data in real-time, such as Hadoop, SPARK, and GRAPHX. Big data analytics using traditional techniques face various challenges, such as fast, accurate and efficient processing of big data in real-time. In addition, the Internet of Things is progressively increasing in various fields, like smart cities, smart homes, and e-health. As the enormous number of connected devices generate huge amounts of data every day, we need sophisticated algorithms to deal, organize, and classify this data in less processing time and space. Similarly, existing techniques and algorithms for deep learning in big data field have several advantages thanks to the two main branches of the deep learning, i.e. convolution and deep belief networks. This book offers insights into these techniques and applications based on these two types of deep learning. Further, it helps students, researchers, and newcomers understand big data analytics based on deep learning approaches. It also discusses various machine learning techniques in concatenation with the deep learning paradigm to support high-end data processing, data classifications, and real-time data processing issues. The classification and presentation are kept quite simple to help the readers and students grasp the basics concepts of various deep learning paradigms and frameworks. It mainly focuses on theory rather than the mathematical background of the deep learning concepts. The book consists of 5 chapters, beginning with an introductory explanation of big data and deep learning techniques, followed by integration of big data and deep learning techniques and lastly the future directions.

This book includes selected papers presented at International Conference on Computational Intelligence, Data Science and Cloud Computing (IEM-ICDC) 2020, organized by the Department of Information Technology, Institute of Engineering & Management, Kolkata, India, during 25-27 September 2020. It presents substantial new research findings about AI and robotics, image processing and NLP, cloud computing and big data analytics as well as in cyber security, blockchain and IoT, and various allied fields. The book serves as a reference resource for researchers and practitioners in academia and industry.

Trends in Deep Learning Methodologies: Algorithms, Applications, and Systems covers deep learning approaches such as neural networks, deep belief networks, recurrent neural networks, convolutional neural networks, deep auto-encoder, and deep generative networks, which have emerged as powerful computational models. Chapters elaborate on these models which have shown significant success in dealing with massive data for a large number of applications, given their capacity to extract complex hidden features and learn efficient representation in

unsupervised settings. Chapters investigate deep learning-based algorithms in a variety of application, including biomedical and health informatics, computer vision, image processing, and more. In recent years, many powerful algorithms have been developed for matching patterns in data and making predictions about future events. The major advantage of deep learning is to process big data analytics for better analysis and self-adaptive algorithms to handle more data. Deep learning methods can deal with multiple levels of representation in which the system learns to abstract higher level representations of raw data. Earlier, it was a common requirement to have a domain expert to develop a specific model for each specific application, however, recent advancements in representation learning algorithms allow researchers across various subject domains to automatically learn the patterns and representation of the given data for the development of specific models. Provides insights into the theory, algorithms, implementation and the application of deep learning techniques Covers a wide range of applications of deep learning across smart healthcare and smart engineering Investigates the development of new models and how they can be exploited to find appropriate solutions

Access real-world documentation and examples for the Spark platform for building large-scale, enterprise-grade machine learning applications. The past decade has seen an astonishing series of advances in machine learning. These breakthroughs are disrupting our everyday life and making an impact across every industry. Next-Generation Machine Learning with Spark provides a gentle introduction to Spark and Spark MLlib and advances to more powerful, third-party machine learning algorithms and libraries beyond what is available in the standard Spark MLlib library. By the end of this book, you will be able to apply your knowledge to real-world use cases through dozens of practical examples and insightful explanations. What You Will Learn Be introduced to machine learning, Spark, and Spark MLlib 2.4.x Achieve lightning-fast gradient boosting on Spark with the XGBoost4J-Spark and LightGBM libraries Detect anomalies with the Isolation Forest algorithm for Spark Use the Spark NLP and Stanford CoreNLP libraries that support multiple languages Optimize your ML workload with the Alluxio in-memory data accelerator for Spark Use GraphX and GraphFrames for Graph Analysis Perform image recognition using convolutional neural networks Utilize the Keras framework and distributed deep learning libraries with Spark Who This Book Is For Data scientists and machine learning engineers who want to take their knowledge to the next level and use Spark and more powerful, next-generation algorithms and libraries beyond what is available in the standard Spark MLlib library; also serves as a primer for aspiring data scientists and engineers who need an introduction to machine learning, Spark, and Spark MLlib.

Deep Learning with PyTorchSimon and Schuster

The five-volume set LNCS 11536, 11537, 11538, 11539, and 11540 constitutes the proceedings of the 19th International Conference on Computational Science, ICCS 2019, held in Faro, Portugal, in June 2019. The total of 65 full papers and 168 workshop papers presented in this book set were carefully reviewed and selected from 573 submissions (228 submissions to the main track and 345 submissions to the workshops). The papers were organized in topical sections named: Part I: ICCS Main Track Part II: ICCS Main Track; Track of Advances in High-Performance Computational Earth Sciences: Applications and Frameworks; Track of Agent-Based Simulations, Adaptive Algorithms and Solvers; Track of Applications of Matrix Methods in Artificial Intelligence and Machine Learning; Track of Architecture, Languages, Compilation and Hardware Support for Emerging and Heterogeneous Systems Part III: Track of Biomedical and Bioinformatics Challenges for Computer Science; Track of Classifier Learning from Difficult Data; Track of Computational Finance and Business Intelligence; Track of Computational Optimization, Modelling and Simulation; Track of Computational Science in IoT and Smart Systems Part IV: Track of Data-Driven Computational Sciences; Track of Machine Learning and Data Assimilation for Dynamical Systems; Track of Marine Computing in the Interconnected World for the Benefit of the Society; Track of Multiscale Modelling and Simulation; Track of Simulations of Flow and Transport: Modeling, Algorithms and Computation Part V: Track of Smart Systems: Computer Vision, Sensor Networks and Machine Learning; Track of Solving Problems with Uncertainties; Track of Teaching Computational Science; Poster Track ICCS 2019 Chapter "Comparing Domain-decomposition Methods for the Parallelization of Distributed Land Surface Models" is available open access under a Creative Commons Attribution 4.0 International License via link.springer.com.

Practical patterns for scaling machine learning from your laptop to a distributed cluster. Scaling up models from standalone devices to large distributed clusters is one of the biggest challenges faced by modern machine learning practitioners. Distributed Machine Learning Patterns teaches you how to scale machine learning models from your laptop to large distributed clusters. In Distributed Machine Learning Patterns, you'll learn how to apply established distributed systems patterns to machine learning projects, and explore new ML-specific patterns as well. Firmly rooted in the real world, this book demonstrates how to apply patterns using examples based in TensorFlow, Kubernetes, Kubeflow, and Argo Workflows. Real-world scenarios, hands-on projects, and clear, practical DevOps techniques let you easily launch, manage, and monitor cloud-native distributed machine learning pipelines. Purchase of the print book includes a free eBook in PDF, Kindle, and ePub formats from Manning Publications.

The two volume set LNCS 12972 + 12973 constitutes the proceedings of the 26th European Symposium on Research in Computer Security, ESORICS 2021, which took place during October 4-8, 2021. The 71 full papers presented in this book were carefully reviewed and selected from 351 submissions. They were organized in topical sections as follows: Part I: network security; attacks; fuzzing; malware; user behavior and underground economy; blockchain; machine learning; automotive; anomaly detection; Part II: encryption; cryptography; privacy; differential privacy; zero knowledge; key exchange; multi-party computation.

This book constitutes the refereed post-conference proceedings of four workshops held at the 15th Asian Conference on Computer Vision, ACCV 2020, which was held in Kyoto, Japan, in November/ December 2020.* The 13 papers were carefully reviewed and selected from the following two workshops: Machine Learning and Computing for Visual Semantic Analysis (MLCSA) and Multi-Visual-Modality Human Activity Understanding (MMHAU). *The conference and workshops were held virtually.

Deep Learning is a subset of Machine Learning where data sets with several layers of complexity can be processed. This book teaches you the different techniques using which deep learning solutions can be implemented at scale, on Apache Spark. This will help you gain experience of implementing your deep learning models in many real-world use cases.

This IBM® Redbooks® publication is a guide about the IBM PowerAI Deep Learning solution. This book provides an introduction to artificial intelligence (AI) and deep learning (DL), IBM PowerAI, and components of IBM PowerAI, deploying IBM PowerAI, guidelines for working with data and creating models, an introduction to IBM SpectrumTM Conductor Deep Learning Impact (DLI), and case scenarios. IBM PowerAI started as a package of software distributions of many of the major DL software frameworks for model training, such as TensorFlow, Caffe, Torch, Theano, and the associated libraries, such as CUDA Deep Neural Network (cuDNN). The IBM PowerAI software is optimized for performance by using the IBM Power SystemsTM servers that are integrated with NVLink. The AI stack foundation starts with servers with accelerators. graphical processing unit (GPU) accelerators are well-suited for the compute-intensive nature of DL training, and servers with the highest CPU to GPU bandwidth, such as IBM Power Systems servers, enable the high-performance data transfer that is required for larger and more complex DL models. This publication targets technical readers, including developers, IT specialists, systems architects, brand specialist, sales team, and anyone looking for a guide about how to understand the IBM PowerAI Deep Learning architecture, framework configuration, application and workload configuration, and user infrastructure.

This book constitutes the refereed proceedings of the Third International Conference on Smart Computing and Communications, SmartCom 2018, held in Tokyo, Japan, in December 2018. The 45 papers presented in this volume were carefully reviewed and selected from 305 submissions. They focus on topics from smart data to smart communications, as well as smart cloud computing to smart security.

This unique text helps make sense of big data in engineering applications using tools and techniques from signal processing. It presents fundamental signal processing theories and software implementations, reviews current research trends and challenges, and describes the techniques used for analysis, design and optimization. Readers will learn about key theoretical issues such as data modelling and representation, scalable and low-complexity information processing and optimization, tensor and sublinear algorithms, and deep learning and software architecture, and their application to a wide range of engineering scenarios. Applications discussed in detail include wireless networking, smart grid systems, and sensor networks and cloud computing. This is the ideal text for researchers and practising engineers wanting to solve practical problems involving large amounts of data, and for students looking to grasp the fundamentals of big data analytics.

The four-volume set LNCS 11334-11337 constitutes the proceedings of the 18th International Conference on Algorithms and Architectures for Parallel Processing, ICA3PP 2018, held in Guangzhou, China, in November 2018. The 141 full and 50 short papers presented were carefully reviewed and selected from numerous submissions. The papers are organized in topical sections on Distributed and Parallel Computing; High Performance Computing; Big Data and Information Processing; Internet of Things and Cloud Computing; and Security and Privacy in Computing.

Implement TensorFlow's offerings such as TensorBoard, TensorFlow.js, TensorFlow Probability, and TensorFlow Lite to build smart automation projects Key Features Use machine learning and deep learning principles to build real-world projects Get to grips with TensorFlow's impressive range of module offerings Implement projects on GANs, reinforcement learning, and capsule network Book Description TensorFlow has transformed the way machine learning is perceived. TensorFlow Machine Learning Projects teaches you how to exploit the benefits—simplicity, efficiency, and flexibility—of using TensorFlow in various real-world projects. With the help of this book, you'll not only learn how to build advanced projects using different datasets but also be able to tackle common challenges using a range of libraries from the TensorFlow ecosystem. To start with, you'll get to grips with using TensorFlow for machine learning projects; you'll explore a wide range of projects using TensorForest and TensorBoard for detecting exoplanets, TensorFlow.js for sentiment analysis, and TensorFlow Lite for digit classification. As you make your way through the book, you'll build projects in various real-world domains, incorporating natural language processing (NLP), the Gaussian process, autoencoders, recommender systems, and Bayesian neural networks, along with trending areas such as Generative Adversarial Networks (GANs), capsule networks, and reinforcement learning. You'll learn how to use the TensorFlow on Spark API and GPU-accelerated computing with TensorFlow to detect objects, followed by how to train and develop a recurrent neural network (RNN) model to generate book scripts. By the end of this book, you'll have gained the required expertise to build full-fledged machine learning projects at work. What you will learn Understand the TensorFlow ecosystem using various datasets and techniques Create recommendation systems for quality product recommendations Build projects using CNNs, NLP, and Bayesian neural networks Play Pac-Man using deep reinforcement learning Deploy scalable TensorFlow-based machine learning systems Generate your own book script using RNNs Who this book is for TensorFlow Machine Learning Projects is for you if you are a data analyst, data scientist, machine learning professional, or deep learning enthusiast with basic knowledge of TensorFlow. This book is also for you if you want to build end-to-end projects in the machine learning domain using supervised, unsupervised, and reinforcement learning techniques

This book provides an insight into ways of inculcating the need for applying mobile edge data analytics in bioinformatics and medicine. The book is a comprehensive reference that provides an overview of the current state of medical treatments and systems and offers emerging solutions for a more personalized approach to the healthcare field. Topics include deep learning methods for applications in object detection and identification, object tracking, human action recognition, and cross-modal and multimodal data analysis. High performance computing systems for applications in healthcare are also discussed. The contributors also include information on microarray data analysis, sequence analysis, genomics based analytics, disease network analysis, and techniques for big data Analytics and health information technology.

The proposed book will discuss various aspects of big data Analytics. It will deliberate upon the tools, technology, applications, use cases and research directions in the field. Chapters would

be contributed by researchers, scientist and practitioners from various reputed universities and organizations for the benefit of readers.
Build next-generation Artificial Intelligence systems with Java Key Features Implement AI techniques to build smart applications using Deeplearning4j Perform big data analytics to derive quality insights using Spark MLlib Create self-learning systems using neural networks, NLP, and reinforcement learning Book Description In this age of big data, companies have larger amount of consumer data than ever before, far more than what the current technologies can ever hope to keep up with. However, Artificial Intelligence closes the gap by moving past human limitations in order to analyze data. With the help of Artificial Intelligence for big data, you will learn to use Machine Learning algorithms such as k-means, SVM, RBF, and regression to perform advanced data analysis. You will understand the current status of Machine and Deep Learning techniques to work on Genetic and Neuro-Fuzzy algorithms. In addition, you will explore how to develop Artificial Intelligence algorithms to learn from data, why they are necessary, and how they can help solve real-world problems. By the end of this book, you'll have learned how to implement various Artificial Intelligence algorithms for your big data systems and integrate them into your product offerings such as reinforcement learning, natural language processing, image recognition, genetic algorithms, and fuzzy logic systems. What you will learn Manage Artificial Intelligence techniques for big data with Java Build smart systems to analyze data for enhanced customer experience Learn to use Artificial Intelligence frameworks for big data Understand complex problems with algorithms and Neuro-Fuzzy systems Design stratagems to leverage data using Machine Learning process Apply Deep Learning techniques to prepare data for modeling Construct models that learn from data using open source tools Analyze big data problems using scalable Machine Learning algorithms Who this book is for This book is for you if you are a data scientist, big data professional, or novice who has basic knowledge of big data and wish to get proficiency in Artificial Intelligence techniques for big data. Some competence in mathematics is an added advantage in the field of elementary linear algebra and calculus.
Due to the growing use of web applications and communication devices, the use of data has increased throughout various industries. It is necessary to develop new techniques for managing data in order to ensure adequate usage. Deep learning, a subset of artificial intelligence and machine learning, has been recognized in various real-world applications such as computer vision, image processing, and pattern recognition. The deep learning approach has opened new opportunities that can make such real-life applications and tasks easier and more efficient. Deep Learning and Neural Networks: Concepts, Methodologies, Tools, and Applications is a vital reference source that trends in data analytics and potential technologies that will facilitate insight in various domains of science, industry, business, and consumer applications. It also explores the latest concepts, algorithms, and techniques of deep learning and data mining and analysis. Highlighting a range of topics such as natural language processing, predictive analytics, and deep neural networks, this multi-volume book is ideally designed for computer engineers, software developers, IT professionals, academicians, researchers, and upper-level students seeking current research on the latest trends in the field of deep learning.
Copyright: 1ad80965c785b128ec9f8d9135b1ae8e