# Learning Real Time Processing With Spark Streaming

This book provides readers with a comprehensive and recent exposition in deep learning and its multidisciplinary applications, with a concentration on advances of deep learning architectures. The book discusses various artificial intelligence (AI) techniques based on deep learning architecture with applications in natural language processing, semantic knowledge, forecasting and many more. The authors shed light on various applications that can benefit from the use of deep learning in pattern recognition, person re-identification in surveillance videos, action recognition in videos, image and video captioning. The book also highlights how deep learning concepts can be interwoven with more modern concepts to yield applications in multidisciplinary fields. Presents a comprehensive look at deep learning and its multidisciplinary applications, concentrating on advances of deep learning architectures; Includes a survey of deep learning problems and solutions, identifying the main open issues, innovations and latest technologies; Shows industrial deep learning in practice with examples/cases, efforts, challenges, and strategic approaches.

This book includes selected papers from the International Conference on Machine Learning and Information Processing (ICMLIP 2019), held at ISB & M School of Technology, Pune, Maharashtra, India, from December 27 to 28, 2019. It presents the latest developments and technical solutions in the areas of advanced computing and data sciences, covering machine learning, artificial intelligence, human-computer interaction, IoT, deep learning, image processing and pattern recognition, and signal and speech processing. Design, process, and analyze large sets of complex data in real time About This Book Get acquainted with transformations and database-level interactions, and ensure the reliability of messages processed using Storm Implement strategies to solve the challenges of real-time data processing Load datasets, build queries, and make recommendations using Spark SQL Who This Book Is For If you are a Big Data architect, developer, or a programmer who wants to develop applications/frameworks to implement real-time analytics using open source technologies, then this book is for you. What You Will Learn Explore big data technologies and frameworks Work through practical challenges and use cases of real-time analytics versus batch analytics Develop real-word use cases for processing and analyzing data in real-time using the programming paradigm of Apache Storm Handle and process real-time transactional data Optimize and tune Apache Storm for varied workloads and production deployments Process and stream data with Amazon Kinesis and Elastic MapReduce Perform interactive and exploratory data analytics using Spark SQL Develop common enterprise architectures/applications for real-time and batch analytics In Detail Enterprise has been striving hard to deal with the challenges of data arriving in real time or near real time. Although there are technologies such as Storm and Spark (and many more) that solve the challenges of real-time data, using the appropriate technology/framework for the right business use case is the key to success. This book provides you with the skills required to quickly design, implement and deploy your real-time analytics using real-world examples of big data use cases. From the beginning of the book, we will cover the basics of varied real-time data processing frameworks and technologies. We will discuss and explain the differences between batch and real-time processing in detail, and will also explore the techniques and programming concepts using Apache Storm. Moving on, we'll familiarize you with "Amazon Kinesis" for real-time data processing on cloud. We will further develop your understanding of real-time analytics through a comprehensive review of Apache Spark along with the high-level architecture and the building blocks of a Spark program. You will learn how to transform your data, get an output from transformations, and persist your results using Spark RDDs, using an interface called Spark SQL to work with Spark. At the end of this book, we will introduce Spark Streaming, the streaming library of Spark, and will walk you through the emerging Lambda Architecture (LA), which provides a hybrid platform for big data processing by combining real-time and precomputed batch data to provide a near real-time view of incoming data. Style and approach This step-by-step is an easy-to-follow, detailed tutorial, filled with practical examples of basic and advanced features. Each topic is explained sequentially and supported by real-world examples and executable code snippets.

Technology driven witty solutions to everyday Managerial Problems Like it is often told "Solutions at your doorstep", we are completely surrounded by profound managerial solutions waiting to be unearthed from our everyday machines in the form of phones, computers, safety devices, automobile etc. The world of machines abounds with managerial thoughts and solutions. This inspiring book provides us with a new approach in problem solving and addresses the diverse challenges faced in managerial functions today. "Learning Management Back From Machines", is the wonderful story of Krish and his latest creation, MANU – an advanced hyper-intelligent, direct-neural interface-capable humanoid, which helps Krish along in deriving managerial solutions from fellow-machines and machine-processes alike. In the process of learning and observing the history of various technological marvels along with the need for these inventions, we discover a whole new dimension of creative intelligence and learning, waiting to reveal itself all over again. The book is aimed at understanding the core essence of how machines have been made to work and help us discover new and innovative solutions to our everyday social and managerial problems. • RELIGIONS TEACH US MANAGEMENT. • STORIES AND FABLES TEACH US MANAGEMENT. • MANAGEMENT THEORIES TEACH US MANAGEMENT. • NOW EVERYDAY MACHINES WILL TEACH US MANAGEMENT

If you are a Java developer who wants to enter into the world of real-time stream processing applications using Apache Storm, then this book is for you. No previous experience in Storm is required as this book starts from the basics. After finishing this book, you will be able to develop not-so-complex Storm applications.

A practical guide to help you tackle different real-time data processing and analytics problems using the best tools for each scenario About This Book Learn about the various challenges in real-time data processing and use the right tools to overcome them This book covers popular tools and frameworks such as Spark, Flink, and Apache Storm to

solve all your distributed processing problems A practical guide filled with examples, tips, and tricks to help you perform efficient Big Data processing in real-time Who This Book Is For If you are a Java developer who would like to be equipped with all the tools required to devise an end-to-end practical solution on real-time data streaming, then this book is for you. Basic knowledge of real-time processing would be helpful, and knowing the fundamentals of Maven, Shell, and Eclipse would be great. What You Will Learn Get an introduction to the established real-time stack Understand the key integration of all the components Get a thorough understanding of the basic building blocks for real-time solution designing Garnish the search and visualization aspects for your real-time solution Get conceptually and practically acquainted with real-time analytics Be well equipped to apply the knowledge and create your own solutions In Detail With the rise of Big Data, there is an increasing need to process large amounts of data continuously, with a shorter turnaround time. Real-time data processing involves continuous input, processing and output of data, with the condition that the time required for processing is as short as possible. This book covers the majority of the existing and evolving open source technology stack for real-time processing and analytics. You will get to know about all the real-time solution aspects, from the source to the presentation to persistence. Through this practical book, you'll be equipped with a clear understanding of how to solve challenges on your own. We'll cover topics such as how to set up components, basic executions, integrations, advanced use cases, alerts, and monitoring. You'll be exposed to the popular tools used in real-time processing today such as Apache Spark, Apache Flink, and Storm. Finally, you will put your knowledge to practical use by implementing all of the techniques in the form of a practical, real-world use case. By the end of this book, you will have a solid understanding of all the aspects of real-time data processing and analytics, and will know how to deploy the solutions in production environments in the best possible manner. Style and Approach In this practical guide to real-time analytics, each chapter begins with a basic high-level concept of the topic, followed by a practical, hands-on implementation of each concept, where you can see the working and execution of it. The book is written in a DIY style, with plenty of practical use cases, well-explained code examples, and relevant screenshots and diagrams.

As metropolises continue to see a growth in population, planners are continually searching for trending methods for utilizing space and seeking the best geographical arrangements for these cities. Professionals have continually used geographic information systems (GIS) to solve these issues; however, limitations in this technology remain prevalent. Integrating multiple-criteria decision analysis and evolutionary computing tools with GIS has created an array of robust solutions for spatial optimization problems in densely populated areas. Interdisciplinary Approaches to Spatial Optimization Issues is a pivotal reference source that provides vital research on advancements within the field of GIS and evolutionary solutions for spatial optimization issues. While highlighting topics such as computing machinery, vehicular routing, and operational research, this publication is ideally designed for practitioners, technicians, developers, academicians, students, government officials, planners, and researchers seeking current research on applications and improvements within spatial optimization and GIS.

Step-by-step guide to different data movement and processing techniques, using Google Cloud Platform Services DESCRIPTION Modern businesses are awash with data, making data-driven decision-making tasks increasingly complex. As a result, relevant technical expertise and analytical skills are required to do such tasks. This book aims to equip you with enough knowledge of Cloud Computing in conjunction with Google Cloud Data platform to succeed in the role of a Cloud data expert. The current market is trending towards the latest cloud technologies, which is the need of the hour. Google being the pioneer, is dominating this space with the right set of cloud services being offered as part of GCP (Google Cloud Platform). At this juncture, this book will be very vital and will cover all the services that are being offered by GCP, putting emphasis on Data services. This book starts with sophisticated knowledge on Cloud Computing. It also explains different types of data services/technology and machine learning algorithm/Pre-Trained API through real-business problems, which are built on the Google Cloud Platform (GCP). With some of the latest business examples and hands-on guide, this book will enable the developers entering the data analytics fields to implement an end-to-end data pipeline, using GCP Data services. Through the course of the book, you will come across multiple industry-wise use cases, like Building Datawarehouse using Big Query, a sample real-time data analytics solution on machine learning and Artificial Intelligence that helped with the business decision, by employing a variety of data science approaches on Google Cloud environment. Whether your business is at the early stage of cloud implementation in its journey or well on its way to digital transformation, Google Cloud's solutions and technologies will always help chart a path to success. This book can be used to develop the GCP concepts in an easy way. It contains many examples showcasing the implementation of a GCP service. It enables the learning of the basic and advance concepts of Google Cloud Data Platform. This book is divided into 7 chapters and provides a detailed description of the core concepts of each of the Data services offered by Google Cloud. KEY FEATURES Learn the basic concept of Cloud Computing along with different Cloud service provides with their supported Models (IaaS/PaaS/SaaS) Learn the basics of Compute Engine, App Engine, Container Engine, Project and Billing setup in the Google Cloud Platform Learn how and when to use Cloud DataFlow, Cloud DataProc and Cloud DataPrep Build real-time data pipeline to support real-time analytics using Pub/Sub messaging service Setting up a fully managed GCP Big Data Cluster using Cloud DataProc for running Apache Spark and Apache Hadoop clusters in a simpler, more cost-efficient manner Learn how to use Cloud Data Studio for visualizing the data on top of Big Query Implement and understand real-world business scenarios for Machine Learning, Data Pipeline Engineering WHAT WILL YOU LEARN By the end of the book, you will have come across different data services and platforms offered by Google Cloud, and how those services/features can be enabled to serve business needs. You will also see a few case studies to put your knowledge to practice and solve business problems such as building a real-time streaming pipeline engine, Scalable Data Warehouse on Cloud, fully managed Hadoop cluster on Cloud and enabling TensorFlow/Machine Learning API's to support real-life business problems. Remember to practice

additional examples to master these techniques. WHO IS THIS BOOK FOR This book is for professionals as well as graduates who want to build a career in Google Cloud data analytics technologies. While no prior knowledge of Cloud Computing or related technologies is assumed, it will be helpful to have some data background and experience. One stop shop for those who wish to get an initial to advance understanding of the GCP data platform. The target audience will be data engineers/professionals who are new, as well as those who are acquainted with the tools and techniques related to cloud and data space. ? Individuals who have basic data understanding (i.e. Data and cloud) and have done some work in the field of data analytics, can refer/use this book to master their knowledge/understanding. ? The highlight of this book is that it will start with the basic cloud computing fundamentals and will move on to cover the advance concepts on GCP cloud data analytics and hence can be referred across multiple different levels of audiences. Table of Contents 1. GCP Overview and Architecture 2. Data Storage in GCP 3. Data Processing in GCP with Pub/Sub and Dataflow 4. Data Processing in GCP with DataPrep and Dataflow 5. Big Query and Data Studio 6. Machine Learning with GCP 7. Sample Use cases and Examples

There is ample evidence that language users, including second-language (L2) users, can predict upcoming information during listening and reading. Yet it is still unclear when, how, and why language users engage in prediction, and what the relation is between prediction and learning. This volume presents a collection of current research, insights, and directions regarding the role of prediction in L2 processing and learning. The contributions in this volume specifically address how different (L1-based) theoretical models of prediction apply to or may be expanded to account for L2 processing, report new insights on factors (linguistic, cognitive, social) that modulate L2 users' engagement in prediction, and discuss the functions that prediction may or may not serve in L2 processing and learning. Taken together, this volume illustrates various fruitful approaches to investigating and accounting for differences in predictive processing within and across individuals, as well as across populations.

There's growing interest in learning how to analyze streaming data in large-scale systems such as web traffic, financial transactions, machine logs, industrial sensors, and many others. But analyzing data streams at scale has been difficult to do well—until now. This practical book delivers a deep introduction to Apache Flink, a highly innovative open source stream processor with a surprising range of capabilities. Authors Ellen Friedman and Kostas Tzoumas show technical and nontechnical readers alike how Flink is engineered to overcome significant tradeoffs that have limited the effectiveness of other approaches to stream processing. You'll also learn how Flink has the ability to handle both stream and batch data processing with one technology. Learn the consequences of not doing streaming well—in retail and marketing, IoT, telecom, and banking and finance Explore how to design data architecture to gain the best advantage from stream processing Get an overview of Flink's capabilities and features, along with examples of how companies use Flink, including in production Take a technical dive into Flink, and learn how it handles time and stateful computation Examine how Flink processes both streaming (unbounded) and batch (bounded) data without sacrificing performance

This book presents cutting-edge research on innovative system interfaces, highlighting both lifecycle development and human–technology interaction, especially in virtual, augmented and mixed reality systems. It describes advanced methodologies and tools for evaluating and improving interface usability, and discusses new models, case studies and good practices. The book addresses the human, hardware, and software factors in the process of developing interfaces for optimizing total system performance, while minimizing costs. It also highlights the forces currently shaping the nature of computing and systems, such as the importance of portability and technologies for reducing power requirements; the need for better assimilation of computation in the environment; and solutions to promote computer and system accessibility for people with special needs. Based on the AHFE 2020 Virtual Conference on Human Factors and Systems Interaction, held on July 16–20, 2020, the book offers a timely survey and a practice-oriented guide for systems interface users and developers alike.

The book describes the emergence of big data technologies and the role of Spark in the entire big data stack. It compares Spark and Hadoop and identifies the shortcomings of Hadoop that have been overcome by Spark. The book mainly focuses on the in-depth architecture of Spark and our understanding of Spark RDDs and how RDD complements big data's immutable nature, and solves it with lazy evaluation, cacheable and type inference. It also addresses advanced topics in Spark, starting with the basics of Scala and the core Spark framework, and exploring Spark data frames, machine learning using Mllib, graph analytics using Graph X and real-time processing with Apache Kafka, AWS Kenisis, and Azure Event Hub. It then goes on to investigate Spark using PySpark and R. Focusing on the current big data stack, the book examines the interaction with current big data tools, with Spark being the core processing layer for all types of data. The book is intended for data engineers and scientists working on massive datasets and big data technologies in the cloud. In addition to industry professionals, it is helpful for aspiring data processing professionals and students working in big data processing and cloud computing environments.

Learning Real Time Processing with Spark Streaming

The book Kafka Streams - Real-time Stream Processing helps you understand the stream processing in general and apply that skill to Kafka streams programming. This book is focusing mainly on the new generation of the Kafka Streams library available in the Apache Kafka 2.x. The primary focus of this book is on Kafka Streams. However, the book also touches on the other Apache Kafka capabilities and concepts that are necessary to grasp the Kafka Streams programming. Who should read this book? Kafka Streams: Real-time Stream Processing is written for software engineers willing to develop a stream processing application using Kafka Streams library. I am also writing this book for data architects and data engineers who are responsible for designing and building the organization's data-centric infrastructure. Another group of people is the managers and architects who do not directly work with Kafka implementation, but they work with the people who implement Kafka Streams at the ground level. What should you already know? This book assumes that the reader is familiar with the basics of Java programming language. The source code and examples in this book are using Java 8, and I will be using Java 8 lambda syntax, so experience with lambda will be helpful. Kafka Streams is a library that runs on Kafka. Having a good fundamental knowledge of Kafka is essential to get the most out of Kafka Streams. I will touch base on the mandatory Kafka concepts for those who are new to Kafka. The book also assumes that you have some familiarity and experience in running and working on the Linux operating system.

Serious games provide a unique opportunity to fully engage students more than traditional teaching approaches. Understanding the best way to utilize these games and the concept of play in an educational setting is imperative for effectual learning in the 21st century. Gamification in Education: Breakthroughs in Research and Practice is an innovative reference source for the latest academic material on the different approaches and issues faced in integrating games within curriculums. Highlighting a range of topics, such as learning through play, virtual worlds, and educational computer games, this publication is ideally designed for educators, administrators, software designers, and stakeholders in all levels of education.

During the 1990s the computing industry has witnessed many advances in mobile and enterprise computing. Many of these advances have been made possible by developments in the areas such as modeling, simulation, and artificial intelligence. Within the different areas of enterprise computing - such as manufacturing, health organisation, and commerce - the need for a disciplined, multifaceted, and

unified approach to modeling and simulation has become essential. This new book provides a forum for scientists, academics, and professionals to present their latest research findings from the various fields: artificial intelligence, collaborative/distributed computing, modeling, and simulation.

Build efficient data flow and machine learning programs with this flexible, multi-functional open-source cluster-computing framework Key Features Master the art of real-time big data processing and machine learning Explore a wide range of use-cases to analyze large data Discover ways to optimize your work by using many features of Spark 2.x and Scala Book Description Apache Spark is an in-memory, cluster-based data processing system that provides a wide range of functionalities such as big data processing, analytics, machine learning, and more. With this Learning Path, you can take your knowledge of Apache Spark to the next level by learning how to expand Spark's functionality and building your own data flow and machine learning programs on this platform. You will work with the different modules in Apache Spark, such as interactive querying with Spark SQL, using DataFrames and datasets, implementing streaming analytics with Spark Streaming, and applying machine learning and deep learning techniques on Spark using MLlib and various external tools. By the end of this elaborately designed Learning Path, you will have all the knowledge you need to master Apache Spark, and build your own big data processing and analytics pipeline quickly and without any hassle. This Learning Path includes content from the following Packt products: Mastering Apache Spark 2.x by Romeo Kienzler Scala and Spark for Big Data Analytics by Md. Rezaul Karim, Sridhar Alla Apache Spark 2.x Machine Learning Cookbook by Siamak Amirghodsi, Meenakshi Rajendran, Broderick Hall, Shuen MeiCookbook What you will learn Get to grips with all the features of Apache Spark 2.x Perform highly optimized real-time big data processing Use ML and DL techniques with Spark MLlib and third-party tools Analyze structured and unstructured data using SparkSQL and GraphX Understand tuning, debugging, and monitoring of big data applications Build scalable and fault-tolerant streaming applications Develop scalable recommendation engines Who this book is for If you are an intermediate-level Spark developer looking to master the advanced capabilities and use-cases of Apache Spark 2.x, this Learning Path is ideal for you. Big data professionals who want to learn how to integrate and use the features of Apache Spark and build a strong big data pipeline will also find this Learning Path useful. To grasp the concepts explained in this Learning Path, you must know the fundamentals of Apache Spark and Scala.

Get started with Apache Flink, the open source framework that powers some of the world's largest stream processing applications. With this practical book, you'll explore the fundamental concepts of parallel stream processing and discover how this technology differs from traditional batch data processing. Longtime Apache Flink committers Fabian Hueske and Vasia Kalavri show you how to implement scalable streaming applications with Flink's DataStream API and continuously run and maintain these applications in operational environments. Stream processing is ideal for many use cases, including low-latency ETL, streaming analytics, and real-time dashboards as well as fraud detection, anomaly detection, and alerting. You can process continuous data of any kind, including user interactions, financial transactions, and IoT data, as soon as you generate them. Learn concepts and challenges of distributed stateful stream processing Explore Flink's system architecture, including its event-time processing mode and fault-tolerance model Understand the fundamentals and building blocks of the DataStream API, including its time-based and statefuloperators Read data from and write data to external systems with exactly-once consistency Deploy and configure Flink clusters Operate continuously running streaming applications

"This comprehensive tutorial will acquaint you with all the aspects of real-time analytics with Apache Spark, one of the trending Big Data processing frameworks on the market today. It will show you how to leverage the features of various components of the Spark framework to efficiently process, analyze, and visualize your data. You will learn how to implement the high velocity streaming operation for data processing in order to perform efficient analytics on your real-time data. You'll analyze data using machine learning techniques and graphs. You'll learn about Spark Streaming and create real-world streaming processing that address all the problems that need to be solved. You'll solve problems using Machine Learning techniques and find out about all the tools available in the MLlibtoolkit. You'll find out how to leverage Graphs to solve real-world problems. At the end of this video, you'll also see some useful Machine Learning algorithms with the help of Spark MLlib and will integrate Spark with R. We'll also make sure you're confident and prepared for graph processing, as you'll learn more about the GraphX API. By the end, you'll be well-versed in the aspects of real-time analytics and implement them with Apache Spark."--Resource description page.

Processing data streams has raised new research challenges over the last few years. This book provides the reader with a comprehensive overview of stream data processing, including famous prototype implementations like the Nile system and the TinyOS operating system. Applications in security, the natural sciences, and education are presented. The huge bibliography offers an excellent starting point for further reading and future research.

This book addresses the major challenges in realizing unmanned aerial vehicles (UAVs) in IoT-based Smart Cities. The challenges tackled vary from cost and energy efficiency to availability and service quality. The aim of this book is to focus on both the design and implementation aspects of the UAV-based approaches in IoT-enabled smart cities' applications that are enabled and supported by wireless sensor networks, 5G, and beyond. The contributors mainly focus on data delivery approaches and their performability aspects. This book is meant for readers of varying disciplines who are interested in implementing the smart planet/environments vision via wireless/wired enabling technologies. Involves the most up to date unmanned aerial vehicles (UAV) assessment and evaluation approaches Includes innovative operational ideas in agriculture, surveillance, rescue, etc. Pertains researchers, scientists, engineers and practitioners in the field of smart cities, IoT, and communications Fadi Al-Turjman received his Ph.D. from Queen's University, Canada. He is a full professor and a research center director at Near East University, Nicosia. He is a leading authority in the area of IoT and intelligent systems. His

publication history spans over 250 publications in addition to his editorialship in top journals such as the IEEE Communication Surveys and Tutorials, and the Elsevier Sustaibable Cities and Society.

Harness the power of Scala to program Spark and analyze tonnes of data in the blink of an eye! About This Book Learn Scala's sophisticated type system that combines Functional Programming and object-oriented concepts Work on a wide array of applications, from simple batch jobs to stream processing and machine learning Explore the most common as well as some complex use-cases to perform large-scale data analysis with Spark Who This Book Is For Anyone who wishes to learn how to perform data analysis by harnessing the power of Spark will find this book extremely useful. No knowledge of Spark or Scala is assumed, although prior programming experience (especially with other JVM languages) will be useful to pick up concepts quicker. What You Will Learn Understand object-oriented & functional programming concepts of Scala In-depth understanding of Scala collection APIs Work with RDD and DataFrame to learn Spark's core abstractions Analysing structured and unstructured data using SparkSQL and GraphX Scalable and fault-tolerant streaming application development using Spark structured streaming Learn machine-learning best practices for classification, regression, dimensionality reduction, and recommendation system to build predictive models with widely used algorithms in Spark MLlib & ML Build clustering models to cluster a vast amount of data Understand tuning, debugging, and monitoring Spark applications Deploy Spark applications on real clusters in Standalone, Mesos, and YARN In Detail Scala has been observing wide adoption over the past few years, especially in the field of data science and analytics. Spark, built on Scala, has gained a lot of recognition and is being used widely in productions. Thus, if you want to leverage the power of Scala and Spark to make sense of big data, this book is for you. The first part introduces you to Scala, helping you understand the object-oriented and functional programming concepts needed for Spark application development. It then moves on to Spark to cover the basic abstractions using RDD and DataFrame. This will help you develop scalable and fault-tolerant streaming applications by analyzing structured and unstructured data using SparkSQL, GraphX, and Spark structured streaming. Finally, the book moves on to some advanced topics, such as monitoring, configuration, debugging, testing, and deployment. You will also learn how to develop Spark applications using SparkR and PySpark APIs, interactive data analytics using Zeppelin, and in-memory data processing with Alluxio. By the end of this book, you will have a thorough understanding of Spark, and you will be able to perform full-stack data analytics with a feel that no amount of data is too big. Style and approach Filled with practical examples and use cases, this book will hot only help you get up and running with Spark, but will also take you farther down the road to becoming a data scientist.

Learn how today's businesses can transform themselves by leveraging real-time data and advanced machine learning analytics. This book provides prescriptive guidance for architects and developers on the design and development of modern Internet of Things (IoT) and Advanced Analytics solutions. In addition, Business in Real-Time Using Azure IoT and Cortana Intelligence Suite offers patterns and practices for those looking to engage their customers and partners through Software-as-a-Service solutions that work on any device. Whether you're working in Health & Life Sciences, Manufacturing, Retail, Smart Cities and Buildings or Process Control, there exists a common platform from which you can create your targeted vertical solutions. Business in Real-Time Using Azure IoT and Cortana Intelligence Suite uses a reference architecture as a road map. Building on Azure's PaaS services, you'll see how a solution architecture unfolds that demonstrates a complete end-to-end IoT and Advanced Analytics scenario. What You'll Learn: Automate your software product life cycle using PowerShell, Azure Resource Manager Templates, and Visual Studio Team Services Implement smart devices using Node.JS and C# Use Azure Streaming Analytics to ingest millions of events Provide both "Hot" and "Cold" path outputs for real-time alerts, data transformations, and aggregation analytics Implement batch processing using Azure Data Factory Create a new form of Actionable Intelligence (AI) to drive mission critical business processes Provide rich Data Visualizations across a wide variety of mobile and web devices Who This Book is For: Solution Architects, Software Developers, Data Architects, Data Scientists, and CIO/CTA Technical Leadership Professionals

Foreword by Oliver Schabenberger, PhD Executive Vice President, Chief Operating Officer and Chief Technology Officer SAS Dive into deep learning! Machine learning and deep learning are ubiquitous in our homes and workplaces—from machine translation to image recognition and predictive analytics to autonomous driving. Deep learning holds the promise of improving many everyday tasks in a variety of disciplines. Much deep learning literature explains the mechanics of deep learning with the goal of implementing cognitive applications fueled by Big Data. This book is different. Written by an expert in high-performance analytics, Deep Learning for Numerical Applications with SAS introduces a new field: Deep Learning for Numerical Applications (DL4NA). Contrary to deep learning, the primary goal of DL4NA is not to learn from data but to dramatically improve the performance of numerical applications by training deep neural networks. Deep Learning for Numerical Applications with SAS presents deep learning concepts in SAS along with step-by-step techniques that allow you to easily reproduce the examples on your high-performance analytics systems. It also discusses the latest hardware innovations that can power your SAS programs: from many-core CPUs to GPUs to FPGAs to ASICs. This book assumes the reader has no prior knowledge of high-performance computing, machine learning, or deep learning. It is intended for SAS developers who want to develop and run the fastest analytics. In addition to discovering the latest trends in hybrid architectures with GPUs and FPGAS, readers will learn how to Use deep learning in SAS Speed up their analytics using deep learning Easily write highly parallel programs using the many task computing paradigms This book is part of the SAS Press program.

Building scalable and fault-tolerant streaming applications made easy with Spark streamingAbout This Book• Process live data streams more efficiently with better fault recovery using Spark Streaming• Implement and deploy real-time log file analysis• Learn about integration with Advance Spark Libraries – GraphX, Spark SQL, and MLib.Who This Book

Is ForThis book is intended for big data developers with basic knowledge of Scala but no knowledge of Spark. It will help you grasp the basics of developing real-time applications with Spark and understand efficient programming of core elements and applications.What You Will Learn• Install and configure Spark and Spark Streaming to execute applications• Explore the architecture and components of Spark and Spark Streaming to use it as a base for other libraries• Process distributed log files in real-time to load data from distributed sources• Apply transformations on streaming data to use its functions• Integrate Apache Spark with the various advance libraries like MLib and GraphX• Apply production deployment scenarios to deploy your applicationIn DetailUsing practical examples with easy-to-follow steps, this book will teach you how to build real-time applications with Spark Streaming.Starting with installing and setting the required environment, you will write and execute your first program for Spark Streaming. This will be followed by exploring the architecture and components of Spark Streaming along with an overview of libraries/functions exposed by Spark. Next you will be taught about various client APIs for coding in Spark by using the use-case of distributed log file processing. You will then apply various functions to transform and enrich streaming data. Next you will learn how to cache and persist datasets. Moving on you will integrate Apache Spark with various other libraries/components of Spark like Mlib, GraphX, and Spark SQL. Finally, you will learn about deploying your application and cover the different scenarios ranging from standalone mode to distributed mode using Mesos, Yarn, and private data centers or on cloud infrastructure.Style and approachA Step-by-Step approach to learn Spark Streaming in a structured manner, with detailed explanation of basic and advance features in an easy-to-follow Style. Each topic is explained sequentially and supported with real world examples and executable code snippets that appeal to the needs of readers with the wide range of experiences.

"This course is your guide to performing real-time data analytics and stream processing with Spark. Use different components and tools such as HDFS, HBase, and Hive to process raw data. Learn how tools such as Hive and Pig aid in this process. In this course, you will start off by learning data analysis techniques with Hadoop using tools such as Hive. Furthermore, you will learn to apply these techniques in real-world big data applications. Also, you will delve into Spark and its related tools to perform real-time data analytics, streaming, and batch processing on your application. Finally, you'll learn how to extend your analytics solutions to the cloud."--Resource description page.

This book introduces the advantages of parallel processing and details how to use it to deal with common signal processing and control algorithms. The text includes examples and end-of-chapter exercises, and case studies to put theoretical concepts into a practical context.

Leverage Scala and Machine Learning to study and construct systems that can learn from data About This Book Explore a broad variety of data processing, machine learning, and genetic algorithms through diagrams, mathematical formulation, and updated source code in Scala Take your expertise in Scala programming to the next level by creating and customizing AI applications Experiment with different techniques and evaluate their benefits and limitations using real-world applications in a tutorial style Who This Book Is For If you're a data scientist or a data analyst with a fundamental knowledge of Scala who wants to learn and implement various Machine learning techniques, this book is for you. All you need is a good understanding of the Scala programming language, a basic knowledge of statistics, a keen interest in Big Data processing, and this book! What You Will Learn Build dynamic workflows for scientific computing Leverage open source libraries to extract patterns from time series Write your own classification, clustering, or evolutionary algorithm Perform relative performance tuning and evaluation of Spark Master probabilistic models for sequential data Experiment with advanced techniques such as regularization and kernelization Dive into neural networks and some deep learning architecture Apply some basic multiarm-bandit algorithms Solve big data problems with Scala parallel collections, Akka actors, and Apache Spark clusters Apply key learning strategies to a technical analysis of financial markets In Detail The discovery of information through data clustering and classification is becoming a key differentiator for competitive organizations. Machine learning applications are everywhere, from self-driving cars, engineering design, logistics, manufacturing, and trading strategies, to detection of genetic anomalies. The book is your one stop guide that introduces you to the functional capabilities of the Scala programming language that are critical to the creation of machine learning algorithms such as dependency injection and implicits. You start by learning data preprocessing and filtering techniques. Following this, you'll move on to unsupervised learning techniques such as clustering and dimension reduction, followed by probabilistic graphical models such as Naive Bayes, hidden Markov models and Monte Carlo inference. Further, it covers the discriminative algorithms such as linear, logistic regression with regularization, kernelization, support vector machines, neural networks, and deep learning. You'll move on to evolutionary computing, multibandit algorithms, and reinforcement learning. Finally, the book includes a comprehensive overview of parallel computing in Scala and Akka followed by a description of Apache Spark and its ML library. With updated codes based on the latest version of Scala and comprehensive examples, this book will ensure that you have more than just a solid fundamental knowledge in machine learning with Scala. Style and approach This book is designed as a tutorial with hands-on exercises using technical analysis of financial markets and corporate data. The approach of each chapter is such that it allows you to understand key concepts easily.

Before you can build analytics tools to gain quick insights, you first need to know how to process data in real time. With this practical guide, developers familiar with Apache Spark will learn how to put this in-memory framework to use for streaming data. You'll discover how Spark enables you to write streaming jobs in almost the same way you write batch jobs. Authors Gerard Maas and François Garillot help you explore the theoretical underpinnings of Apache Spark. This comprehensive guide features two sections that compare and contrast the streaming APIs Spark now supports: the original Spark Streaming library and the newer Structured Streaming API. Learn fundamental stream processing concepts and examine different streaming architectures Explore Structured Streaming through practical examples; learn different aspects of stream processing in

detail Create and operate streaming jobs and applications with Spark Streaming; integrate Spark Streaming with other Spark APIs Learn advanced Spark Streaming techniques, including approximation algorithms and machine learning algorithms Compare Apache Spark to other stream processing projects, including Apache Storm, Apache Flink, and Apache Kafka Streams

This book offers a comprehensive overview of ensemble learning in the field of feature selection (FS), which consists of combining the output of multiple methods to obtain better results than any single method. It reviews various techniques for combining partial results, measuring diversity and evaluating ensemble performance. With the advent of Big Data, feature selection (FS) has become more necessary than ever to achieve dimensionality reduction. With so many methods available, it is difficult to choose the most appropriate one for a given setting, thus making the ensemble paradigm an interesting alternative. The authors first focus on the foundations of ensemble learning and classical approaches, before diving into the specific aspects of ensembles for FS, such as combining partial results, measuring diversity and evaluating ensemble performance. Lastly, the book shows examples of successful applications of ensembles for FS and introduces the new challenges that researchers now face. As such, the book offers a valuable guide for all practitioners, researchers and graduate students in the areas of machine learning and data mining.

Learn how to apply the principles of machine learning to time series modeling with this indispensable resource Machine Learning for Time Series Forecasting with Python is an incisive and straightforward examination of one of the most crucial elements of decision-making in finance, marketing, education, and healthcare: time series modeling. Despite the centrality of time series forecasting, few business analysts are familiar with the power or utility of applying machine learning to time series modeling. Author Francesca Lazzeri, a distinguished machine learning scientist and economist, corrects that deficiency by providing readers with comprehensive and approachable explanation and treatment of the application of machine learning to time series forecasting. Written for readers who have little to no experience in time series forecasting or machine learning, the book comprehensively covers all the topics necessary to: Understand time series forecasting concepts, such as stationarity, horizon, trend, and seasonality Prepare time series data for modeling Evaluate time series forecasting models' performance and accuracy Understand when to use neural networks instead of traditional time series models in time series forecasting Machine Learning for Time Series Forecasting with Python is full real-world examples, resources and concrete strategies to help readers explore and transform data and develop usable, practical time series forecasts. Perfect for entry-level data scientists, business analysts, developers, and researchers, this book is an invaluable and indispensable guide to the fundamental and advanced concepts of machine learning applied to time series modeling.

Real-time or applied digital signal processing courses are offered as follow-ups to conventional or theory-oriented digital signal processing courses in many engineering programs for the purpose of teaching students the technical know-how for putting signal processing algorithms or theory into practical use. These courses normally involve access to a teaching laboratory that is equipped with hardware boards, in particular DSP boards, together with their supporting software. A number of textbooks have been written discussing how to achieve real-time implementation on these hardware boards. This book discusses how to use smartphones as hardware boards for real-time implementation of signal processing algorithms as an alternative to the hardware boards that are used in signal processing laboratory courses. The fact that mobile devices, in particular smartphones, have become powerful processing platforms led to the development of this book enabling students to use their own smartphones to run signal processing algorithms in real-time considering that these days nearly all students possess smartphones. Changing the hardware platforms that are currently used in applied or real-time signal processing courses to smartphones creates a truly mobile laboratory experience or environment for students. In addition, it relieves the cost burden associated with using dedicated signal processing boards noting that the software development tools for smartphones are free of charge and are well-maintained by smartphone manufacturers. This book is written in such a way that it can be used as a textbook for real-time or applied digital signal processing courses offered at many universities. Ten lab experiments that are commonly encountered in such courses are covered in the book. This book is written primarily for those who are already familiar with signal processing concepts and are interested in their real-time and practical aspects. Similar to existing real-time courses, knowledge of C programming is assumed. This book can also be used as a self-study guide for those who wish to become familiar with signal processing app development on either Android or iPhone smartphones.

Crosslinguistic Influence and Second Language Learning provides a comprehensive overview of what is currently known about prior language knowledge and experience in second language learning. Three bodies of research are critically reviewed to achieve this goal: (i) theories of language learning that attribute critical roles to prior experience in explaining second language development, (ii) empirical studies of second language learning that have investigated roles for crosslinguistic influence, and (iii) instructional studies that have supported second language learning by addressing the negative effects of crosslinguistic influence. Using this foundation, new research directions and theorization in the field of second language acquisition are proposed. This book will serve as an excellent resource for students and scholars with interests in (instructed) second language learning, applied linguistics, cognitive psychology, psycholinguistics, and language education.

Immersive technology as an umbrella concept consists of multiple emerging technologies including augmented reality (AR), virtual reality (VR), gaming, simulation, and 3D printing. Research has shown immersive technology provides unique learning opportunities for experiential learning, multiple perspectives, and knowledge transfer. Due to its role in influencing learners? cognitive and affective processes, it is shown to have great potential in changing the educational landscape in the decades to come. However, there is a lack of general cognitive and affective theoretical framework to guide the diverse aspects of immersive technology research. In fact, lacking the cognitive and affective theoretical framework has begun to hamper the design and application of immersive technology in schools and related professional training. Cognitive and Affective Perspectives on Immersive Technology in Education is an essential research book that explores methods and implications for the design and implementation of upcoming immersive technologies in pedagogical and professional development settings. The book includes case studies that highlight the cognitive and affective processes in immersive technology as well

as the successful applications of immersive technology in education. Featuring a wide range of topics such as curriculum design, K-12 education, and mobile learning, this book is ideal for academicians, educators, policymakers, curriculum developers, instructional designers, administrators, researchers, and students.

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.

Grid technologies are rising with the next generation of Internet by defining a powerful computing paradigm. Grid could be used as a technology 'glue' providing users with a uniform way to access resources by means of several devices. These technologies can provide a support for Technology Enhanced Learning (TEL) by enabling new learning environments based on collaboration, real direct experience, personalization, ubiquity, accessibility and contextualization. Nevertheless, to be effectively used in TEL, Grid must be complemented with other elements like semantics and educational modelling; leading to the concept of 'Learning Grid' as defined in the homonymous Special Interest Group (SIG) of the European Network of Excellence 'Kaleidoscope: Shaping the Scientific Evolution of Technology Enhanced Learning'. The key challenge that Kaleidoscope is facing is the scientific and structural integration of the European TEL research. In this context, the Learning Grid SIG aims at contributing to the achievement of an improvement in TEL practices through the definition of open, distributed and pervasive environments for effective human learning taking into account that effective learning requires an active attitude of learners and that learning is a social and collaborative activity so requiring a technology that allows for active and realistic experiments, personalization, knowledge creation and evolution, as well as autonomous and dynamic creation of communities. The first section of the book is about the concept of Learning Grid and related technologies. The second chapter analyses and compares existing languages for the dynamic composition of distributed learning resources and services in a Learning Grid.

Artificial intelligence and its various components are rapidly engulfing almost every professional industry. Specific features of AI that have proven to be vital solutions to numerous real-world issues are machine learning and deep learning. These intelligent agents unlock higher levels of performance and efficiency, creating a wide span of industrial applications. However, there is a lack of research on the specific uses of machine/deep learning in the professional realm. Machine Learning and Deep Learning in Real-Time Applications provides emerging research exploring the theoretical and practical aspects of machine learning and deep learning and their implementations as well as their ability to solve real-world problems within several professional disciplines including healthcare, business, and computer science. Featuring coverage on a broad range of topics such as image processing, medical improvements, and smart grids, this book is ideally designed for researchers, academicians, scientists, industry experts, scholars, IT professionals, engineers, and students seeking current research on the multifaceted uses and implementations of machine learning and deep learning across the globe.

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.

With the widespread interest in digital entertainment and the advances in the technologies of computer graphics, multimedia and virtual reality technologies, the new area of "Edutainment" has been accepted as a union of education and computer entertainment. Edutainment is recognized as an effective way of learning through a medium, such as a computer, software, games or AR/VR applications, that both educates and entertains. The Edutainment conference series was established and followed as a special event for the new interests in e-learning and digital entertainment. The main purpose of Edutainment conferences is the discussion, presentation, and information exchange of scientific and technological developments in the new community. The Edutainment conference series is a very interesting opportunity for researchers, engineers, and graduate students who wish to communicate at these international annual events. The conference series includes plenary invited talks, workshops, tutorials, paper presen- tion tracks, and panel discussions. The Edutainment conference series was initiated in Hangzhou, China in 2006. Following the success of the first (Edutainment 2006 in Hangzhou, China), the second (Edutainment 2007 in Hong Kong, China), and the third events (Edutainment 2008 in Nanjing, China), Edutainment 2009 was held August 9–11, 2009 in Banff, Canada. This year, we received 116

submissions from 25 different countries and regions - cluding Austria, Canada, China, Denmark, Finland, France, Germany, Greece, Hong Kong, Italy, Japan, Korea, Malaysia, Mexico, The Netherlands, Norway, Portugal, Singapore, Spain, Sweden, Switzerland, Taiwan, Trinidad and Tobago, UK, and USA.

The groundbreaking Encyclopedia of Ecology provides an authoritative and comprehensive coverage of the complete field of ecology, from general to applied. It includes over 500 detailed entries, structured to provide the user with complete coverage of the core knowledge, accessed as intuitively as possible, and heavily cross-referenced. Written by an international team of leading experts, this revolutionary encyclopedia will serve as a one-stop-shop to concise, stand-alone articles to be used as a point of entry for undergraduate students, or as a tool for active researchers looking for the latest information in the field. Entries cover a range of topics, including: Behavioral Ecology Ecological Processes Ecological Modeling Ecological Engineering Ecological Indicators Ecological Informatics Ecosystems Ecotoxicology Evolutionary Ecology General Ecology Global Ecology Human Ecology System Ecology The first reference work to cover all aspects of ecology, from basic to applied Over 500 concise, stand-alone articles are written by prominent leaders in the field Article text is supported by full-color photos, drawings, tables, and other visual material Fully indexed and cross referenced with detailed references for further study Writing level is suited to both the expert and non-expert Available electronically on ScienceDirect shortly upon publication