Cook
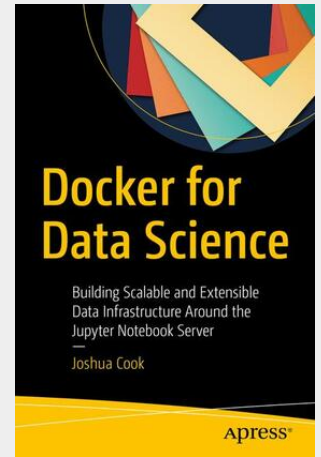
Docker for Data Science

Building Scalable and Extensible Data Infrastructure Around the Jupyter Notebook Server

Learn Docker "infrastructure as code" technology to define a system for performing standard but non-trivial data tasks on medium- to large-scale data sets, using Jupyter as the master controller. It is not uncommon for a real-world data set to fail to be easily managed. The set may not fit well into access memory or may require prohibitively long processing. These are significant challenges to skilled software engineers and they can render the standard Jupyter system unusable. As a solution to this problem, Docker for Data Science proposes using Docker. You will learn how to use existing pre-compiled public images created by the major open-source technologies—Python, Jupyter, Postgres—as well as using the Dockerfile to extend these images to suit your specific purposes. The Docker-Compose technology is examined and you will learn how it can be used to build a linked system with Python churning data behind the scenesand Jupyter managing these background tasks. Best practices in using existing images are explored as well as developing your own images to deploy state-of-the-art machine learning and optimization algorithms. What You'll Learn - Master interactive development using the Jupyter platform - Run and build Docker containers from scratch and from publicly available open-source images - Write infrastructure as code using the docker-compose tool and its docker-compose.yml file type - Deploy a multi-service data science application across a cloud-based system Who This Book Is For Data scientists, machine learning engineers, artificial intelligence researchers, Kagglers, and software developers

Learn Docker "infrastructure as code" technology to define a system for performing standard but non-trivial data tasks on medium- to large-scale data sets, using Jupyter as the master controller. It is not uncommon for a real-world data set to fail to be easily managed. The set may not fit well into access memory or may require prohibitively long processing. These are significant challenges to skilled software engineers and they can render the standard Jupyter system unusable. As a solution to this problem, Docker for Data Science proposes using Docker. You will learn how to use existing pre-compiled public images created by the major open-source technologies—Python, Jupyter, Postgres—as well as using the Dockerfile to extend these images to suit your specific purposes. The Docker-Compose technology is examined and you will learn how it can be used to build a linked system with Python churning data behind the scenesand Jupyter managing these background tasks. Best practices in using existing images are explored as well as developing your own images to deploy state-of-the-art machine learning and optimization algorithms. What You'll Learn: - Master interactive development using the Jupyter platform - Run and build Docker containers from scratch and from publicly available open-source images - Write infrastructure as code using the docker-compose tool and its docker-compose.yml file type - Deploy a multi-service data science application across a cloud-based system

64,19 €
59,99 € (zzgl. MwSt.)

Lieferfrist: bis zu 10 Tage

Artikelnummer: 9781484230114
Medium: Buch
ISBN: 978-1-4842-3011-4
Verlag: Apress
Erscheinungstermin: 25.08.2017
Sprache(n): Englisch
Auflage: 1. Auflage 2017
Produktform: Kartoniert
Gewicht: 4336 g
Seiten: 257
Format (B x H): 155 x 235 mm

21.07.2024 | 22:27 Uhr