

1. Record Nr.	UNINA9910483986303321
Autore	Gupta Pramod
Titolo	Introduction to machine learning in the cloud with Python : concepts and practices // Pramod Gupta, Naresh K. Sehgal
Pubbl/distr/stampa	Cham, Switzerland : , : Springer, , [2021] ©2021
ISBN	3-030-71270-2
Descrizione fisica	1 online resource (293 pages)
Disciplina	006.31
Soggetti	Machine learning
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Intro -- Foreword -- Preface -- About the Book -- Contents -- Acronyms -- Part I: Concepts -- Chapter 1: Machine Learning Concepts -- 1.1 Terminology -- 1.2 What Is Machine Learning? -- 1.2.1 Mitchell's Notion of Machine Learning -- 1.3 What Does Learning Mean for a Computer? -- 1.4 Difference Between ML and Traditional Programming -- 1.5 How Do Machines Learn? -- 1.6 Steps to Apply ML -- 1.7 Paradigms of Learning -- 1.7.1 Supervised Machine Learning -- 1.7.2 Unsupervised Machine Learning -- 1.7.3 Reinforcement Machine Learning -- 1.7.3.1 Types of Problems in Machine Learning -- 1.8 Machine Learning in Practice -- 1.9 Why Use Machine Learning? -- 1.10 Why Machine Learning Now? -- 1.11 Classical Tasks for Machine Learning -- 1.12 Applications of Machine Learning -- 1.12.1 Applications in Our Daily Life -- 1.13 ML Computing Needs -- 1.14 Machine Learning in the Cloud -- 1.15 Tools Used in Machine Learning -- 1.16 Points to Ponder -- References -- Chapter 2: Machine Learning Algorithms -- 2.1 Why Choose Machine Learning? -- 2.2 Supervised Machine Learning Algorithms -- 2.2.1 Regression -- 2.2.2 Classification -- 2.2.3 Machine Learning Algorithms: Supervised Learning -- 2.2.4 Machine Learning Algorithms: Unsupervised Learning -- 2.2.4.1 Clustering -- 2.2.4.2 Dimension Reduction -- 2.2.4.3 Anomaly Detection -- 2.2.5 Machine Learning Algorithms That Use Unsupervised Learning -- 2.3 Considerations in Choosing an Algorithm -- 2.4 What Are the Most Common and Popular Machine Learning

Algorithms? -- 2.4.1 Linear Regression -- 2.4.2 Two Types of Linear Regression -- 2.4.2.1 Simple Linear Regression -- 2.4.2.2 Multiple Linear Regression -- 2.4.2.3 Assumptions of Linear Regression -- 2.4.2.4 Advantages -- 2.4.2.5 Disadvantages -- 2.4.2.6 Sample Python Code for Linear Regression -- 2.4.3 K-Nearest Neighbors (KNN) -- 2.4.3.1 Assumptions. 2.4.3.2 How Does KNN Algorithm Works? -- 2.4.3.3 Advantages -- Sample Python Code for KNN Algorithm -- 2.4.4 Logistic Regression -- 2.4.4.1 Types of Logistic Regression -- 2.4.4.2 Assumptions -- 2.4.4.3 Advantages -- 2.4.4.4 Disadvantages -- Sample Python Code for Implementing Logistic Regression -- 2.4.5 Naive Bayes Classifier Algorithm -- 2.4.5.1 Additive Smoothing -- 2.4.5.2 Types of Naive Bayes Model -- 2.4.5.3 Assumptions -- 2.4.5.4 How Naive Bayes Algorithm Works? -- 2.4.5.5 Advantages -- 2.4.5.6 Disadvantages -- Sample Python Code for Naive Bayes Model -- 2.4.6 Support Vector Machine Algorithm -- 2.4.6.1 Assumptions -- 2.4.6.2 Types of SVM -- 2.4.6.3 Advantages -- 2.4.6.4 Disadvantages -- Sample Python Code for SVM -- 2.4.7 Decision Trees -- 2.4.7.1 Information Gain -- 2.4.7.2 Gini Index -- 2.4.7.3 Decision Tree Terminology -- 2.4.7.4 Assumptions -- 2.4.7.5 How Does the Decision Tree Classifier Work? -- 2.4.7.6 Advantages -- 2.4.7.7 Disadvantages -- Sample Python Code for Decision Tree -- 2.4.8 Ensemble Learning -- 2.4.8.1 Types of Ensemble Learning -- Boosting -- Bootstrap Aggregation (Bagging) -- 2.4.9 Random Forests -- 2.4.9.1 Assumptions -- 2.4.9.2 How Does Random Forest Algorithm Work? -- 2.4.9.3 Advantages -- 2.4.9.4 Disadvantages -- Sample Python Code for Random Forest -- 2.4.10 K-Means Clustering Algorithm -- 2.4.10.1 Assumptions -- 2.4.10.2 How Does K-Means Algorithm Work? -- 2.4.10.3 Convergence Criterion -- 2.4.10.4 Advantages -- 2.4.10.5 Disadvantages -- Sample Python Code for K-Means -- 2.4.11 Artificial Neural Networks -- 2.4.11.1 Advantages -- 2.4.11.2 Disadvantages -- 2.5 Usage of ML Algorithms -- 2.6 Performance Metrics of ML Algorithms -- 2.6.1 Testing Data -- 2.6.2 Performance Metrics for Classification Models -- 2.6.2.1 Confusion Matrix -- 2.6.2.2 Regression Metrics -- 2.7 Most Popular Machine Learning Software Tools. 2.8 Machine Learning Platforms -- 2.8.1 Alteryx Analytics -- 2.8.2 H2O.ai -- 2.8.3 KNIME Analytics Platform -- 2.8.4 RapidMiner -- 2.8.5 Databricks Unified Analytics Platform -- 2.8.6 Microsoft's Azure Machine Learning Studio -- 2.8.7 Google's Analytics Platform -- 2.8.8 IBM Watson -- 2.8.9 Amazon Web Services (AWS) -- 2.9 Points to Ponder -- References -- Chapter 3: Deep Learning and Cloud Computing -- 3.1 Deep Learning (DL) -- 3.2 Historical Trends -- 3.3 How Do Deep Learning Algorithm Learn? -- 3.3.1 Activation Functions -- 3.4 Architectures -- 3.4.1 Deep Neural Network (DNN) -- 3.4.2 Recurrent Neural Network (RNN) -- 3.4.3 Convolutional Neural Networks (CNN) -- 3.5 Choosing a Network -- 3.6 Deep Learning Development Flow -- 3.7 What Is Deep About Deep Learning? -- 3.8 Data Used for Deep Learning -- 3.9 Difference Between Machine Learning and Deep Learning -- 3.10 Why Deep Learning Became Popular Now? -- 3.11 Should You Always Use Deep Learning Instead of Machine Learning? -- 3.12 Why Is Deep Learning Important? -- 3.13 What Are the Drawbacks of Deep Learning? -- 3.14 Which Deep Learning Software Frameworks Are Available? -- 3.15 Classical Problems of Deep Learning Solves -- 3.15.1 Image Classification -- 3.15.2 Natural Language Processing -- 3.16 The Future of Deep Learning -- 3.17 Points to Ponder -- References -- Chapter 4: Cloud Computing Concepts -- 4.1 Roots of Cloud Computing -- 4.2 Key Characteristics of Cloud Computing -- 4.3 Various Cloud Stakeholders

-- 4.4 Pain Points in Cloud Computing -- 4.5 AI and ML in Cloud --
4.6 Expanding Cloud Reach -- 4.7 Future Trends -- 4.8 Summary --
4.9 Points to Ponder -- References -- Part II: Practices -- Chapter 5:
Practical Aspects in Machine Learning -- 5.1 Preprocessing Data -- 5.2
Challenges in Data Preparation -- 5.3 When to Use Data Preprocessing?
-- 5.4 Framework for Data Preparation Techniques.
5.4.1 Data Preparation -- 5.4.2 Data Selection (Feature Selection) --
5.4.3 Data Preprocessing -- 5.4.4 Data Cleaning -- 5.4.5 Insufficient
Data -- 5.4.6 Non-representative Data -- 5.4.7 Substandard Data --
5.4.8 Data Transformation -- 5.4.9 Handling Missing Values -- 5.5
Modification of Categorical or Text Values to Numerical Values -- 5.6
Feature Scaling -- 5.6.1 Techniques of Feature Scaling -- 5.6.1.1
Feature Scaling: Standardization -- 5.6.1.2 Feature Scaling:
Normalization (Min-Max Normalization) -- 5.7 Inconsistent Values --
5.8 Duplicated Values -- 5.9 Feature Aggregation -- 5.10 Feature
Sampling -- 5.10.1 Sampling Without Replacement -- 5.10.2 Sampling
with Replacement -- 5.11 Multicollinearity and Its Impact -- 5.12
Feature Selection -- 5.12.1 Importance of Feature Selection -- 5.12.2
How Many Features to Have in the Model? -- 5.12.3 Types of Feature
Selection -- 5.12.3.1 Filter Method -- 5.12.3.2 Wrapper Methods --
5.12.3.3 Embedded Methods (Fig. 5.7) -- LASSO Regression -- Ridge
Regression -- 5.13 Dimensionality Reduction -- 5.13.1 Principal
Component Analysis (PCA) -- 5.13.2 Linear Discriminant Analysis --
5.13.3 t-Distributed Stochastic Neighbor Embedding (t-SNE) -- 5.14
Dealing with Imbalanced Data -- 5.14.1 Use the Right Evaluation
Metrics -- 5.14.2 Sampling-Based Approaches -- 5.14.3 Algorithm
Based Approach -- 5.15 Points to Ponder -- References -- Chapter 6:
Information Security and Cloud Computing -- 6.1 Information Security
Background and Context -- 6.2 Privacy Issues -- 6.3 Security Concerns
of Cloud Operating Models -- 6.4 Secure Transmissions, Storage, and
Computation -- 6.5 A Few Key Challenges Related to Cloud Computing
and Virtualization -- 6.6 Security Practices for Cloud Computing -- 6.7
Role of ML for Cybersecurity -- 6.8 Summary -- 6.9 Points to Ponder
-- References -- Chapter 7: Examples of Analytics in the Cloud.
7.1 Background -- 7.2 Analytics Services in the Cloud -- 7.3
Introduction to MapReduce -- 7.4 Introduction to Hadoop -- 7.5
Examples of Cloud-Based ML -- 7.5.1 Cloud Security Monitoring Using
AWS -- 7.5.2 Greener Energy Future with ML in GCP -- 7.5.3 Monorail
Monitoring in Azure -- 7.5.4 Detecting Online Hate Speech Using NLP
-- 7.6 Future Possibilities -- 7.7 Summary -- 7.8 Points to Ponder --
References -- Chapter 8: Health Care in the Cloud: A Few Case Studies
-- 8.1 Introduction -- 8.2 Existing TCD Solution -- 8.3 Trail of Bubbles
-- 8.4 Moving Data to the Cloud -- 8.5 A Reader in the Cloud -- 8.6
Cloud-Based Collaborative Tools -- 8.7 Multi-Cloud Solutions -- 8.8
UCSD Antibigram: Using Unclassifiable Data -- 8.9 Next Steps -- 8.10
Summary -- 8.11 Points to Ponder -- References -- Chapter 9: Trends
in Hardware-Based AI and ML -- 9.1 Revisiting the History of AI -- 9.2
Current Limitations of AI and ML -- 9.3 Emergence of AI Hardware
Accelerators -- 9.3.1 Use of GPUs -- 9.3.2 Use of FPGAs -- 9.3.3
Dedicated AI Accelerators Using ASICs -- 9.4 Cerebras's Wafer Scale AI
Engine -- 9.5 Google Cloud TPUs -- 9.6 Amazon's Inference Engine --
9.7 Intel's Movidius VPU -- 9.8 Apple's AI Ecosystem -- 9.9 Summary
-- 9.10 Points to Ponder -- References -- Appendix A -- AI/ML for
App Store Predictions -- Using Python for App Metrics Predictions in
Google Play Store -- Step 1: Obtain Data -- Step 2: Scrub and Prepare
the Data -- Step 2.1: Null Value Removal/Fills -- Step 2.2: Duplicate
Removal -- Exact Duplicates (876) -- Slight Variations (294) -- Step
2.3: Data Type Conversions -- Step 2.4: Dummy Variables -- Step 2.5:

Rescale -- Step 3: Data Exploration -- Step 4: Categorization -- Step
4.1: Google App Ratings -- Step 4.2: Google App Reviews -- Step 4.3:
Google App Sizes -- Step 4. 4: Google App Installs -- Step 4.5: Type --
Step 4.6: Price.
Step 4.7: Content Rating.
