1. Record Nr. UNISA996499867703316

Autore	El Morr Christo <1966->
Titolo	Machine learning for practical decision making : a multidisciplinary perspective with applications from healthcare, engineering and business analytics / / Christo El Morr [and three others]
Pubbl/distr/stampa	Cham, Switzerland : , : Springer, , [2022] ©2022
ISBN	3-031-16990-5
Descrizione fisica	1 online resource (475 pages)
Collana	International series in operations research & management science ; ; Volume 334
Disciplina	658.403
Soggetti	Decision making - Data processing Machine learning Presa de decisions Processament de dades Aprenentatge automàtic Llibres electrònics
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di bibliografia	Includes bibliographical references and index.
Nota di contenuto	Intro -- Preface -- Contents -- Chapter 1: Introduction to Machine Learning -- 1.1 Introduction to Machine Learning -- 1.2 Origin of Machine Learning -- 1.3 Growth of Machine Learning -- 1.4 How Machine Learning Works -- 1.5 Machine Learning Building Blocks -- 1.5.1 Data Management and Exploration -- 1.5.1.1 Data, Information, and Knowledge -- 1.5.1.2 Big Data -- 1.5.1.3 OLAP Versus OLTP -- 1.5.1.4 Databases, Data Warehouses, and Data Marts -- 1.5.1.5 Multidimensional Analysis Techniques -- 1.5.1.5.1 Slicing and Dicing -- 1.5.1.5.2 Pivoting -- 1.5.1.5.3 Drill-Down, Roll-Up, and Drill-Across -- 1.5.2 The Analytics Landscape -- 1.5.2.1 Types of Analytics (Descriptive, Diagnostic, Predictive, Prescriptive) -- 1.5.2.1.1 Descriptive Analytics -- 1.5.2.1.2 Diagnostic Analytics -- 1.5.2.1.3 Predictive Analytics -- 1.5.2.1.4 Prescriptive Analytics -- 1.6 Conclusion -- 1.7 Key Terms -- 1.8 Test Your Understanding -- 1.9 Read More -- 1.10 Lab -- 1.10.1 Introduction to R -- 1.10.2

Introduction to RStudio -- 1.10.2.1 RStudio Download and Installation -- 1.10.2.2 Install a Package -- 1.10.2.3 Activate Package -- 1.10.2.4 User Readr to Load Data -- 1.10.2.5 Run a Function -- 1.10.2.6 Save Status -- 1.10.3 Introduction to Python and Jupyter Notebook IDE -- 1.10.3.1 Python Download and Installation -- 1.10.3.2 Jupyter Download and Installation -- 1.10.3.3 Load Data and Plot It Visually -- 1.10.3.4 Save the Execution -- 1.10.3.5 Load a Saved Execution -- 1.10.3.6 Upload a Jupyter Notebook File -- 1.10.4 Do It Yourself -- References -- Chapter 2: Statistics -- 2.1 Overview of the Chapter -- 2.2 Definition of General Terms -- 2.3 Types of Variables -- 2.3.1 Measures of Central Tendency -- 2.3.1.1 Measures of Dispersion -- 2.4 Inferential Statistics -- 2.4.1 Data Distribution -- 2.4.2 Hypothesis Testing -- 2.4.3 Type I and II Errors.
2.4.4 Steps for Performing Hypothesis Testing -- 2.4.5 Test Statistics -- 2.4.5.1 Student´s t-test -- 2.4.5.2 One-Way Analysis of Variance -- 2.4.5.3 Chi-Square Statistic -- 2.4.5.4 Correlation -- 2.4.5.5 Simple Linear Regression -- 2.5 Conclusion -- 2.6 Key Terms -- 2.7 Test Your Understanding -- 2.8 Read More -- 2.9 Lab -- 2.9.1 Working Example in R -- 2.9.1.1 Statistical Measures Overview -- 2.9.1.2 Central Tendency Measures in R -- 2.9.1.3 Dispersion in R -- 2.9.1.4 Statistical Test Using p-value in R -- 2.9.2 Working Example in Python -- 2.9.2.1 Central Tendency Measure in Python -- 2.9.2.2 Dispersion Measures in Python -- 2.9.2.3 Statistical Testing Using p-value in Python -- 2.9.3 Do It Yourself -- 2.9.4 Do More Yourself (Links to Available Datasets for Use) -- References -- Chapter 3: Overview of Machine Learning Algorithms -- 3.1 Introduction -- 3.2 Data Mining -- 3.3 Analytics and Machine Learning -- 3.3.1 Terminology Used in Machine Learning -- 3.3.2 Machine Learning Algorithms: A Classification -- 3.4 Supervised Learning -- 3.4.1 Multivariate Regression -- 3.4.1.1 Multiple Linear Regression -- 3.4.1.2 Multiple Logistic Regression -- 3.4.2 Decision Trees -- 3.4.3 Artificial Neural Networks -- 3.4.3.1 Perceptron -- 3.4.4 Naïve Bayes Classifier -- 3.4.5 Random Forest -- 3.4.6 Support Vector Machines (SVM) -- 3.5 Unsupervised Learning -- 3.5.1 K-Means -- 3.5.2 K-Nearest Neighbors (KNN) -- 3.5.3 AdaBoost -- 3.6 Applications of Machine Learning -- 3.6.1 Machine Learning Demand Forecasting and Supply Chain Performance [42] -- 3.6.2 A Case Study on Cervical Pain Assessment with Motion Capture [43] -- 3.6.3 Predicting Bank Insolvencies Using Machine Learning Techniques [44] -- 3.6.4 Deep Learning with Convolutional Neural Network for Objective Skill Evaluation in Robot-Assisted Surgery [45] -- 3.7 Conclusion -- 3.8 Key Terms.
3.9 Test Your Understanding -- 3.10 Read More -- 3.11 Lab -- 3.11.1 Machine Learning Overview in R -- 3.11.1.1 Caret Package -- 3.11.1.2 ggplot2 Package -- 3.11.1.3 mlBench Package -- 3.11.1.4 Class Package -- 3.11.1.5 DataExplorer Package -- 3.11.1.6 Dplyr Package -- 3.11.1.7 KernLab Package -- 3.11.1.8 Mlr3 Package -- 3.11.1.9 Plotly Package -- 3.11.1.10 Rpart Package -- 3.11.2 Supervised Learning Overview -- 3.11.2.1 KNN Diamonds Example -- 3.11.2.1.1 Loading KNN Algorithm Package -- 3.11.2.1.2 Loading Dataset for KNN -- 3.11.2.1.3 Preprocessing Data -- 3.11.2.1.4 Scaling Data -- 3.11.2.1.5 Splitting Data and Applying KNN Algorithm -- 3.11.2.1.6 Model Performance -- 3.11.3 Unsupervised Learning Overview -- 3.11.3.1 Loading K-Means Clustering Package -- 3.11.3.2 Loading Dataset for K-Means Clustering Algorithm -- 3.11.3.3 Preprocessing Data -- 3.11.3.4 Executing K-Means Clustering Algorithm -- 3.11.3.5 Results Discussion -- 3.11.4 Python Scikit-Learn Package Overview -- 3.11.5 Python Supervised Learning Machine (SML) -- 3.11.5.1 Using Scikit-Learn Package -- 3.11.5.2 Loading Diamonds Dataset Using

Python -- 3.11.5.3 Preprocessing Data -- 3.11.5.4 Splitting Data and Executing Linear Regression Algorithm -- 3.11.5.5 Model Performance Explanation -- 3.11.5.6 Classification Performance -- 3.11.6 Unsupervised Machine Learning (UML) -- 3.11.6.1 Loading Dataset for Hierarchical Clustering Algorithm -- 3.11.6.2 Running Hierarchical Algorithm and Plotting Data -- 3.11.7 Do It Yourself -- 3.11.8 Do More Yourself -- References -- Chapter 4: Data Preprocessing -- 4.1 The Problem -- 4.2 Data Preprocessing Steps -- 4.2.1 Data Collection -- 4.2.2 Data Profiling, Discovery, and Access -- 4.2.3 Data Cleansing and Validation -- 4.2.4 Data Structuring -- 4.2.5 Feature Selection -- 4.2.6 Data Transformation and Enrichment.
4.2.7 Data Validation, Storage, and Publishing -- 4.3 Feature Engineering -- 4.3.1 Feature Creation -- 4.3.2 Transformation -- 4.3.3 Feature Extraction -- 4.4 Feature Engineering Techniques -- 4.4.1 Imputation -- 4.4.1.1 Numerical Imputation -- 4.4.1.2 Categorical Imputation -- 4.4.2 Discretizing Numerical Features -- 4.4.3 Converting Categorical Discrete Features to Numeric (Binarization) -- 4.4.4 Log Transformation -- 4.4.5 One-Hot Encoding -- 4.4.6 Scaling -- 4.4.6.1 Normalization (Min-Max Normalization) -- 4.4.6.2 Standardization (Z-Score Normalization) -- 4.4.7 Reduce the Features Dimensionality -- 4.5 Overfitting -- 4.6 Underfitting -- 4.7 Model Selection: Selecting the Best Performing Model of an Algorithm -- 4.7.1 Model Selection Using the Holdout Method -- 4.7.2 Model Selection Using Cross-Validation -- 4.7.3 Evaluating Model Performance in Python -- 4.8 Data Quality -- 4.9 Key Terms -- 4.10 Test Your Understanding -- 4.11 Read More -- 4.12 Lab -- 4.12.1 Working Example in Python -- 4.12.1.1 Read the Dataset -- 4.12.1.2 Split the Dataset -- 4.12.1.3 Impute Data -- 4.12.1.4 One-Hot-Encode Data -- 4.12.1.5 Scale Numeric Data: Standardization -- 4.12.1.6 Create Pipelines -- 4.12.1.7 Creating Models -- 4.12.1.8 Cross-Validation -- 4.12.1.9 Hyperparameter Finetuning -- 4.12.2 Working Example in Weka -- 4.12.2.1 Missing Values -- 4.12.2.2 Discretization (or Binning) -- 4.12.2.3 Data Normalization and Standardization -- 4.12.2.4 One-Hot-Encoding (Nominal to Numeric) -- 4.12.3 Do It Yourself -- 4.12.3.1 Lenses Dataset -- 4.12.3.2 Nested Cross-Validation -- 4.12.4 Do More Yourself -- References -- Chapter 5: Data Visualization -- 5.1 Introduction -- 5.2 Presentation and Visualization of Information -- 5.2.1 A Taxonomy of Graphs -- 5.2.2 Relationships and Graphs -- 5.2.3 Dashboards -- 5.2.4 Infographics -- 5.3 Building Effective Visualizations.
5.4 Data Visualization Software -- 5.5 Conclusion -- 5.6 Key Terms -- 5.7 Test Your Understanding -- 5.8 Read More -- 5.9 Lab -- 5.9.1 Working Example in Tableau -- 5.9.1.1 Getting a Student Copy of Tableau Desktop -- 5.9.1.2 Learning with Tableau´s how-to Videos and Resources -- 5.9.2 Do It Yourself -- 5.9.2.1 Assignment 1: Introduction to Tableau -- 5.9.2.2 Assignment 2: Data Manipulation and Basic Charts with Tableau -- 5.9.3 Do More Yourself -- 5.9.3.1 Assignment 3: Charts and Dashboards with Tableau -- 5.9.3.2 Assignment 4: Analytics with Tableau -- References -- Chapter 6: Linear Regression -- 6.1 The Problem -- 6.2 A Practical Example -- 6.3 The Algorithm -- 6.3.1 Modeling the Linear Regression -- 6.3.2 Gradient Descent -- 6.3.3 Gradient Descent Example -- 6.3.4 Batch Versus Stochastic Gradient Descent -- 6.3.5 Examples of Error Functions -- 6.3.6 Gradient Descent Types -- 6.3.6.1 Stochastic Gradient Descent -- 6.3.6.2 Batch Gradient -- 6.4 Final Notes: Advantages, Disadvantages, and Best Practices -- 6.5 Key Terms -- 6.6 Test Your Understanding -- 6.7 Read More -- 6.8 Lab -- 6.8.1 Working Example in R -- 6.8.1.1 Load Diabetes Dataset -- 6.8.1.2 Preprocess

Diabetes Dataset -- 6.8.1.3 Choose Dependent and Independent Variables -- 6.8.1.4 Visualize Your Dataset -- 6.8.1.5 Split Data into Test and Train Datasets -- 6.8.1.6 Create Linear Regression Model and Visualize it -- 6.8.1.7 Calculate Confusion Matrix -- 6.8.1.8 Gradient Descent -- 6.8.2 Working Example in Python -- 6.8.2.1 Load USA House Prices Dataset -- 6.8.2.2 Explore Housing Prices Visually -- 6.8.2.3 Preprocess Data -- 6.8.2.4 Split Data and Scale Features -- 6.8.2.5 Create and Visualize Model Using the LinearRegression Algorithm -- 6.8.2.6 Evaluate Performance of LRM -- 6.8.2.7 Optimize LRM Manually with Gradient Descent.
6.8.2.8 Create and Visualize a Model Using the Stochastic Gradient Descent (SGD).