

1. Record Nr.	UNINA9910768168903321
Autore	Acito Frank
Titolo	Predictive Analytics with KNIME : Analytics for Citizen Data Scientists // Frank Acito
Pubbl/distr/stampa	Cham, Switzerland : , : Springer, , [2023] ©2023
ISBN	3-031-45630-0
Edizione	[First edition.]
Descrizione fisica	1 online resource (317 pages)
Disciplina	001.42
Soggetti	Predictive analytics
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 Analytics -- 1.1 The Growing Emphasis on Analytics -- 1.2 Applications of Analytics -- 1.3 The Citizen Data Scientist -- 1.4 The Analytics Process -- 1.5 Summary -- References -- Chapter 2: Problem Definition -- 2.1 Expert Views on Problem Definition -- 2.2 A Telecom Problem Definition Example -- 2.3 Defining the Analytics Problem -- 2.4 Structured Versus Unstructured Problems -- 2.5 Getting Started with Defining the Problem -- 2.6 Summary -- References -- Chapter 3: Introduction to KNIME -- 3.1 KNIME Features -- 3.2 The KNIME Workbench -- 3.3 Learning to Use KNIME -- 3.4 KNIME Extensions and Integrations -- 3.5 Data Types in KNIME -- 3.6 Example: Predicting Heart Disease with KNIME -- 3.7 Example: Preparation of Hospital Data Using KNIME -- 3.8 Flow Variables -- 3.9 Loops in KNIME -- 3.10 Metanodes and Components in KNIME -- 3.11 Summary -- Appendices -- Appendix 1: Integrating R into KNIME -- Appendix 2: Regex for Search Patterns -- Chapter 4: Data Preparation -- 4.1 Obtaining the Needed Data -- 4.2 Data Cleaning -- 4.3 Data Cleaning Nodes in KNIME -- 4.4 Missing Values -- 4.5 Dealing with Missing Values -- 4.6 Outliers -- 4.7 Feature Engineering -- 4.8 Example of Using KNIME for Data Preparation -- 4.9 Summary -- References -- Chapter 5: Dimensionality Reduction -- 5.1 Problems with Large Numbers of Variables -- 5.2 Approaches to Dimension Reduction -- 5.3 Principal Components Analysis -- 5.4 Example of Using PCA -- 5.5 Intuition

and Algebra behind Principal Components -- 5.6 Summary --  
References -- Chapter 6: Ordinary Least Squares Regression -- 6.1  
Intuition About Simple Linear Regression -- 6.2 Multiple Regression --  
6.3 Building a Predictive Regression Model -- Variable Types  
in Regression -- Coding Nominal Variables -- Coding Ordinal Variables  
-- 6.4 Nonlinear Relationships.  
Regression with a Nonlinear Relationship -- Using a Polynomial Model  
-- Transformations to Deal with Nonlinearity -- 6.5 Evaluating  
Predictive Accuracy -- 6.6 Example Applications of Regression --  
Predicting Home Prices -- Comparison of OLS and Stepwise Regression  
-- Using Regularization for Predictor Selection -- Regularization  
Formula -- Comparisons of Prediction Accuracy -- 6.7 Summary --  
References -- Chapter 7: Logistic Regression -- 7.1 Intuition About  
Binary Logistic Regression -- 7.2 Modeling Probabilities -- 7.3  
Estimating Logistic Regression Parameters -- 7.4 Example Using  
Simulated Data -- The Logistic Analysis Results -- Estimating  
Probabilities -- 7.5 The Nonlinearity of Logistic Regression Coefficients  
-- Change in  $X_1$  (0.07 to +0.03) with  $X_2 = 1.50$  -- Change in  $X_1$  (+0.  
170 to +0.270) with  $X_2 = 1.50$  -- Change in  $X_1$  (0.07 to +0.03)  
with  $X_2 = 2.00$  -- Change in  $X_1$  (0.07 to +0.03) with  $X_2 = 2.00$  --  
Changes in Odds, Log Odds, and Probability -- 7.6 Interpreting Logistic  
Results Using Log Odds -- 7.7 Evaluating Classification Models --  
Confusion Matrices -- Which Metrics to Use with Confusion Matrices?  
-- ROC Curves -- Metrics with Multiple-Level Categorical Targets --  
7.8 Example: Predicting Employee Retention Using Logistic Regression  
-- Results for the Analysis of Employee Turnover -- 7.9 Predictor  
Interpretation and Importance -- An Approximate Method for Predictor  
Interpretation -- 7.10 Example: Predicting Heart Disease Using Logistic  
Regression -- Results for Heart Disease Prediction -- 7.11 Regularized  
Logistic Regression -- Applying Regularization to the Heart Disease  
Data -- Interpreting the Coefficients -- 7.12 Asymmetric Benefits  
and Costs -- 7.13 Multinomial Logistic Regression -- 7.14 Summary --  
Appendix: Cohen's Kappa -- References -- Chapter 8: Classification  
and Regression Trees -- 8.1 Classification Trees.  
8.2 Applications of Decision Trees -- 8.3 Developing Classification  
Trees -- 8.4 Growing Decision Trees Using Gini Impurity --  
Demonstrating Gini Calculations -- 8.5 Pruning to Avoid Overfitting --  
Pre-pruning -- Post-pruning -- Reduced Error Pruning -- Cost-  
Complexity Pruning -- Pruning Using the Minimum Description Length  
Principle -- Recommendations on Pruning -- 8.6 Missing Values  
in Decision Tree Analyses -- Ignoring Missing Data -- Imputation  
Techniques -- Using Machine Learning -- 8.7 Outliers in Classification  
Trees -- 8.8 Predicting Churn with a Classification Tree -- 8.9  
Regression Trees -- 8.10 Example: Head Acceleration in a Simulated  
Motorcycle Accident -- 8.11 Strengths and Weaknesses of Decision  
Trees -- Strengths of Decision Trees -- Weaknesses of Decision Trees  
-- 8.12 Summary -- References -- Chapter 9: Naïve Bayes -- 9.1 A  
Thought Problem -- Analysis of the Thought Problem -- 9.2 Bayes'  
Theorem Illustrated -- Updating the Probabilities -- What Happens  
with More Predictors? -- 9.3 Illustration of Naïve Bayes with a "Toy"  
Data Set -- Calculations Needed for the Naïve Bayes' Model -- Results  
of Probability Calculations -- 9.4 The Assumption of Conditional  
Independence -- 9.5 Naïve Bayes with Continuous Predictors -- 9.6  
Laplace Smoothing -- 9.7 Example of Naïve Bayes Applied to the Heart  
Disease Data -- Heart Disease Predictions with Naïve Bayes -- 9.8  
Example of Naïve Bayes Applied to Detecting Spam -- Accuracy  
of SPAM/HAM Detection -- 9.9 Summary and Comments on Naïve  
Bayes -- References -- Chapter 10: k Nearest Neighbors -- 10.1 How

kNN Works -- 10.2 A Two-Dimensional Graphic Example of kNN -- 10.3 Example Application of kNN to Diagnosing Heart Disease -- 10.4 kNN for Continuous Targets -- 10.5 kNN for Multiclass Target Variables -- 10.6 Summary -- References -- Chapter 11: Neural Networks.

11.1 What Are Artificial Neural Networks? -- 11.2 The Learning Process for Artificial Neural Networks -- 11.3 Example of a Single-Layer Artificial Neuron -- 11.4 Example of a Multilayer Perceptron -- 11.5 Example Application of a Multilayer Perceptron with Multi-level Categorical Target -- 11.6 Considerations for Using Neural Nets -- 11.7 Example: Using Neural Nets to Predict Credit Status -- 11.8 Example: Neural Nets to Predict Used Car Prices -- 11.9 Summary -- References -- Chapter 12: Ensemble Models -- 12.1 Creating Ensemble Models -- 12.2 Ensemble Models Based on Decision Trees -- 12.3 Example of Ensemble Modeling with a Continuous Target -- 12.4 Example of Ensemble Modeling for a Binary Target -- 12.5 Summary -- References -- Chapter 13: Cluster Analysis -- 13.1 How Many Clusters Are There? -- 13.2 Recommended Steps in Running a Cluster Analysis -- 13.3 Hierarchical Cluster Analysis -- 13.4 k-Means Clustering -- 13.5 Density-Based Clustering -- 13.6 Fuzzy Cluster Analysis -- 13.7 Cluster Validation -- 13.8 Summary -- References -- Chapter 14: Communication and Deployment -- 14.1 Writing and Presenting the Final Report -- 14.2 Data Visualization -- 14.3 Deploying Predictive Models -- 14.4 Summary -- References -- Index.

---