

1. Record Nr.	UNINA9910861099003321
Autore	Michelucci Umberto
Titolo	Fundamental Mathematical Concepts for Machine Learning in Science
Pubbl/distr/stampa	Cham : , : Springer International Publishing AG, , 2024 ©2024
ISBN	3-031-56431-6
Edizione	[1st ed.]
Descrizione fisica	1 online resource (259 pages)
Lingua di pubblicazione	Inglese
Formato	Materiale a stampa
Livello bibliografico	Monografia
Nota di contenuto	Intro -- Preface -- Acknowledgements -- Contents -- Acronyms -- Chapter 1 Introduction -- 1.1 Introduction -- 1.2 Choice of Topics -- 1.3 Prerequisites -- 1.4 Book Structure -- 1.5 About This Book -- 1.6 Warnings, Info and Examples -- 1.7 Optional and Advanced Material -- 1.8 Further Exploration and Reading -- 1.9 References -- 1.10 Let us Start -- References -- Chapter 2 Machine Learning: History and Terminology -- 2.1 Brief History of Machine Learning -- 2.2 Machine Learning in Science -- 2.3 Types of Machine Learning -- References -- Chapter 3 Calculus and Optimisation for Machine Learning -- 3.1 Motivation -- 3.2 Concept of Limit -- 3.3 Derivative and its Properties -- 3.4 Partial Derivative -- 3.5 Gradient -- 3.6 Extrema of a Function -- 3.7 Optimisation for Machine Learning -- 3.8 Introduction to Optimisation for Neural Networks -- 3.8.1 First Definition of Learning -- 3.8.2 Constrained vs. Unconstrained Optimisation -- 3.9 Optimization Algorithms -- 3.9.1 Line Search and Trust Region Approaches -- 3.9.2 Steepest Descent -- 3.9.3 Additional Directions for the Line Search Approach -- 3.9.4 The Gradient Descent Algorithm -- 3.9.5 Choosing the Right Learning Rate -- 3.9.6 Variations of Gradient Descent -- 3.9.6.1 Mini-batch Gradient Descent -- 3.9.6.2 Stochastic Gradient Descent -- 3.9.7 How to Choose the Right Mini-batch Size -- 3.9.8 Stochastic Gradient Descent and Fractals -- 3.10 Conclusions -- References -- Chapter 4 Linear Algebra -- 4.1 Motivation -- 4.2 Vectors -- 4.2.1 Geometrical Interpretation of Vectors -- 4.2.2 Norm of Vectors -- 4.2.3 Dot Product -- 4.2.4 Cross

Product -- 4.3 Matrices -- 4.3.1 Sum, Subtraction and Transpose -- 4.3.2 Multiplication of Matrices and Vectors -- 4.3.3 Inverse and Trace -- 4.3.4 Determinant -- 4.3.5 Matrix Calculus and Linear Regression -- 4.4 Relevance for Machine Learning. 4.5 Eigenvectors and Eigenvalues -- 4.6 Principal Component Analysis -- 4.6.1 Basis of a Vector Space -- 4.6.2 Definition of a Vector Space -- 4.6.3 Linear Transformations (maps) -- 4.6.4 PCA Formalisation -- 4.6.5 Covariance Matrix -- 4.6.6 Overview of Assumptions -- 4.6.7 PCA with Eigenvectors and Eigenvalues -- 4.6.8 One Implementation Limitation -- References -- Chapter 5 Statistics and Probability for Machine Learning -- 5.1 Motivation -- 5.2 Random Experiments and Variables -- 5.3 Algebra of Sets -- 5.4 Probability -- 5.4.1 Relative Frequency Interpretation of Probability -- 5.4.2 Probability as a Set Function -- 5.4.3 Axiomatic Definition of Probability -- 5.4.4 Properties of Probability Functions -- 5.5 The Softmax Function -- 5.5.1 Softmax Range of Applications -- 5.6 Some Theorems about Probability Functions -- 5.7 Conditional Probability -- 5.8 Bayes Theorem -- 5.9 Bayes Error -- 5.10 Naïve Bayes Classifier -- 5.11 Distribution Functions -- 5.11.1 Cumulative Distribution Function (CDF) -- 5.11.2 Probability Density (PDF) and Mass Functions (PMF) -- 5.12 Expected Values and its Properties -- 5.13 Variance and its Properties -- 5.13.1 Properties -- 5.14 Normal Distribution -- 5.15 Other Distributions -- 5.16 The MSE and its Distribution -- 5.16.1 Moment Generating Functions -- 5.16.2 Central Limit Theorem -- 5.17 Central Limit Theorem without Mathematics -- References -- Chapter 6 Sampling Theory (a.k.a. Creating a Dataset Properly) -- 6.1 Introduction -- 6.2 Research Questions and Hypotheses -- 6.2.1 Research Questions -- 6.2.2 Hypothesis -- 6.2.3 Relevance of Hypothesis and Research Questions in Machine Learning -- 6.3 Survey Populations -- 6.4 Survey Samples -- 6.4.1 Non-probability Sampling -- 6.4.2 Probability Sampling -- 6.5 Stratification and Clustering -- 6.6 Random Sampling without Replacement -- 6.7 Random Sampling with Replacement. 6.8 Bootstrapping -- 6.9 Random Stratified Sampling -- 6.10 Sampling in Machine Learning -- References -- Chapter 7 Model Validation and Selection -- 7.1 Introduction -- 7.2 Bias-Variance Tradeoff -- 7.3 Bias-Variance Tradeoff - a Mathematical Discussion -- 7.4 High-Variance Low-Bias regime -- 7.5 Low-Variance High-Bias regime -- 7.6 Overfitting and Underfitting -- 7.7 The Simple Split Approach (a.k.a. Hold-out Approach) -- 7.8 Data Leakage -- 7.8.1 Data Leakage with Correlated Observations -- 7.8.2 Stratified Sampling -- 7.9 Monte Carlo Cross-Validation -- 7.10 Monte-Carlo Cross Validation with Bootstrapping -- 7.11 k-Fold Cross Validation -- 7.12 The Leave-One-Out Approach -- 7.13 Choosing the Cross-Validation Approach -- 7.14 Model Selection -- 7.14.1 Model Selection with Supervised Learning -- 7.14.2 Model Selection with Unsupervised Learning -- 7.15 Qualitative Criteria for Model Selection -- References -- Chapter 8 Unbalanced Datasets and Machine Learning Metrics -- 8.1 Introduction -- 8.2 A Simple Example -- 8.3 Approaches to Deal with Unbalanced Datasets -- 8.3.1 Oversampling -- 8.3.2 (Random) Undersampling -- 8.4 Synthetic Minority Oversampling TEchnique (SMOTE) -- 8.5 Summary of Methods for Dealing with Unbalanced Datasets -- 8.6 Important Metrics -- 8.6.1 The Notion of Metric -- 8.6.1.1 .The MSE is a Metric -- 8.6.1.2 . 1 is a Metric -- 8.6.2 Confusion Matrix -- 8.6.3 Sensitivity and Specificity -- 8.6.4 Precision -- 8.6.5 -score -- 8.6.6 Balanced Accuracy -- 8.6.7 Receiving Operating Characteristic (ROC) Curve -- 8.6.8 Probability Interpretation of the AUC -- References -- Chapter 9 Hyper-parameter Tuning -- 9.1 Introduction -- 9.2 Black-box

Optimisation -- 9.2.1 Notes on Black-box Functions -- 9.3 The Problem of Hyper-parameter Tuning -- 9.4 Sample Black-box Problem -- 9.4.1 Grid Search -- 9.4.2 Random Search. 9.4.3 Coarse to Fine Optimisation -- 9.4.4 Sampling on a Logarithmic Scale -- 9.5 Overview of Approaches for Hyper-parameter Tuning -- References -- Chapter 10 Feature Importance and Selection -- 10.1 Introduction -- 10.2 Feature Importance Taxonomy -- 10.2.1 Filter Methods -- 10.2.2 Wrapper Methods -- 10.2.3 Embedded Methods -- 10.3 Forward Feature Selection -- 10.3.1 Forward Feature Selection Practical Example -- 10.4 Backward Feature Elimination -- 10.5 Permutation Feature Importance -- 10.6 Information Content Elimination -- 10.7 Summary -- 10.8 SHapley Additive exPlanations (SHAP) -- References -- Index.
