

User Guide

1. Supervised learning

- ▶ [1.1. Linear Models](#)
- ▶ [1.2. Linear and Quadratic Discriminant Analysis](#)
- [1.3. Kernel ridge regression](#)
- ▶ [1.4. Support Vector Machines](#)
- ▶ [1.5. Stochastic Gradient Descent](#)
- ▶ [1.6. Nearest Neighbors](#)
- ▶ [1.7. Gaussian Processes](#)
- ▶ [1.8. Cross decomposition](#)
- ▶ [1.9. Naive Bayes](#)
- ▶ [1.10. Decision Trees](#)
- ▶ [1.11. Ensemble methods](#)
- ▶ [1.12. Multiclass and multioutput algorithms](#)
- ▶ [1.13. Feature selection](#)
- ▶ [1.14. Semi-supervised learning](#)
- [1.15. Isotonic regression](#)
- ▶ [1.16. Probability calibration](#)
- ▶ [1.17. Neural network models \(supervised\)](#)

2. Unsupervised learning

- ▶ [2.1. Gaussian mixture models](#)
- ▶ [2.2. Manifold learning](#)
- ▶ [2.3. Clustering](#)
- ▶ [2.4. Biclustering](#)
- ▶ [2.5. Decomposing signals in components \(matrix factorization problems\)](#)
- ▶ [2.6. Covariance estimation](#)
- ▶ [2.7. Novelty and Outlier Detection](#)

Toggle Menu

- ▶ [2.8. Density Estimation](#)
- ▶ [2.9. Neural network models \(unsupervised\)](#)

3. Model selection and evaluation

- ▶ [3.1. Cross-validation: evaluating estimator performance](#)
- ▶ [3.2. Tuning the hyper-parameters of an estimator](#)
- ▶ [3.3. Metrics and scoring: quantifying the quality of predictions](#)
- ▶ [3.4. Validation curves: plotting scores to evaluate models](#)

4. Inspection

- ▶ [4.1. Partial Dependence and Individual Conditional Expectation plots](#)
- ▶ [4.2. Permutation feature importance](#)

5. Visualizations

- ▶ [5.1. Available Plotting Utilities](#)

6. Dataset transformations

- ▶ [6.1. Pipelines and composite estimators](#)
- ▶ [6.2. Feature extraction](#)
- ▶ [6.3. Preprocessing data](#)
- ▶ [6.4. Imputation of missing values](#)
- ▶ [6.5. Unsupervised dimensionality reduction](#)
- ▶ [6.6. Random Projection](#)
- ▶ [6.7. Kernel Approximation](#)
- ▶ [6.8. Pairwise metrics, Affinities and Kernels](#)
- ▶ [6.9. Transforming the prediction target \(y\).](#)

7. Dataset loading utilities

- ▶ [7.1. Toy datasets](#)
- ▶ [7.2. Real world datasets](#)

- ▶ [7.4. Loading other datasets](#)

8. Computing with scikit-learn

- ▶ [8.1. Strategies to scale computationally: bigger data](#)
- ▶ [8.2. Computational Performance](#)
- ▶ [8.3. Parallelism, resource management, and configuration](#)

9. Model persistence

- ▶ [9.1. Python specific serialization](#)
- [9.2. Interoperable formats](#)

10. Common pitfalls and recommended practices

- [10.1. Inconsistent preprocessing](#)
- ▶ [10.2. Data leakage](#)
- ▶ [10.3. Controlling randomness](#)

© 2007 - 2021, scikit-learn developers (BSD License). [Show this page source](#)