Unsupervised dimensionality reduction

If your number of features is high, it may be useful to reduce it with an unsupervised step prior to supervised steps. Many of the :ref:`unsupervised-learning` methods implement a `transform` method that can be used to reduce the dimensionality. Below we discuss two specific example of this pattern that are heavily used.

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scikit-learn-main\doc\modules\[scikit-learn-main][doc][modules]unsupervised_reduction.rst, line 8); *backlink*

Unknown interpreted text role "ref".

Pipelining

The unsupervised data reduction and the supervised estimator can be chained in one step. See :ref:`pipeline`.

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scikit-learn-main\doc\modules\[scikit-learn-main][doc][modules]unsupervised_reduction.rst, line 16); *backlink*

Unknown interpreted text role "ref".

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scikit-learn-main\doc\modules\[scikit-learn-main][doc][modules]unsupervised_reduction.rst, line 19)

Unknown directive type "currentmodule".

```
.. currentmodule:: sklearn
```

PCA: principal component analysis

:class:`decomposition.PCA` looks for a combination of features that capture well the variance of the original features. See :ref:`decompositions`.

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scikit-learn-main\doc\modules\[scikit-learn-main][doc][modules]unsupervised_reduction.rst, line 24); *backlink*

Unknown interpreted text role "class".

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scikit-learn-main\doc\modules\[scikit-learn-main][doc][modules]unsupervised_reduction.rst, line 24); *backlink*

Unknown interpreted text role "ref".

Examples

- :ref:`sphx_glr_auto_examples_applications_plot_face_recognition.py`

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scikit-learn-main\doc\modules\[scikit-learn-main][doc][modules]unsupervised_reduction.rst, line 29); *backlink*

Unknown interpreted text role "ref".

Random projections

The module: :mod:`random_projection` provides several tools for data reduction by random projections. See the relevant section of the documentation: :ref:`random_projection`.

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scikit-

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scikit-learn-main\doc\modules\[scikit-learn-main][doc][modules]unsupervised_reduction.rst, line 34); *backlink*

Unknown interpreted text role "mod".

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scikit-learn-main\doc\modules\[scikit-learn-main][doc][modules]unsupervised_reduction.rst, line 34); *backlink*

Unknown interpreted text role "ref".

Examples

- :ref:`sphx_glr_auto_examples_miscellaneous_plot_johnson_lindenstrauss_bound.py`

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scikit-learn-main\doc\modules\[scikit-learn-main][doc][modules]unsupervised_reduction.rst, line 40); *backlink*

Unknown interpreted text role "ref".

Feature agglomeration

:class:`cluster.FeatureAgglomeration` applies :ref:`hierarchical_clustering` to group together features that behave similarly.

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scikit-learn-main\doc\modules\[scikit-learn-main][doc][modules]unsupervised_reduction.rst, line 45); *backlink*

Unknown interpreted text role "class".

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scikit-learn-main\doc\modules\[scikit-learn-main][doc][modules]unsupervised_reduction.rst, line 45); *backlink*

Unknown interpreted text role "ref".

Examples

- :ref:`sphx_glr_auto_examples_cluster_plot_feature_agglomeration_vs_univariate_selection.py`

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scikit-learn-main\doc\modules\[scikit-learn-main][doc][modules]unsupervised_reduction.rst, line 51); *backlink*

Unknown interpreted text role "ref".

- :ref:`sphx_glr_auto_examples_cluster_plot_digits_agglomeration.py`

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scikit-learn-main\doc\modules\[scikit-learn-main][doc][modules]unsupervised_reduction.rst, line 52); *backlink*

Unknown interpreted text role "ref".

Feature scaling

Note that if features have very different scaling or statistical properties, :class:`cluster.FeatureAgglomeration` may not be able to capture the links between related features. Using a :class:`preprocessing.StandardScaler` can be useful in these settings.

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scikit-learn-main\doc\modules\[scikit-learn-main][doc][modules]unsupervised_reduction.rst, line 56); *backlink*

Unknown interpreted text role "class".

System Message: ERROR/3 (D:\onboarding-resources\sample-onboarding-resources\scikit-learn-main\doc\modules\[scikit-learn-main][doc][modules]unsupervised_reduction.rst, line 56); *backlink*

Unknown interpreted text role "class".