11.1. Array API support (experimental) ¶

The Array API specification defines a standard API for all array manipulation libraries with a NumPy-like API.

Some scikit-learn estimators that primarily rely on NumPy (as opposed to using Cython) to implement the algorithmic logic of their `fit`, `predict` or `transform` methods can be configured to accept any Array API compatible input datastructures and automatically dispatch operations to the underlying namespace instead of relying on NumPy.

At this stage, this support is **considered experimental** and must be enabled explicitly as explained in the following.

Note: Currently, only `cupy.array_api` and `numpy.array_api` are known to work with scikit-learn's estimators.

11.1.1. Example usage

Here is an example code snippet to demonstrate how to use CuPy to run `LinearDiscriminantAnalysis` on a GPU:

```
>>> from sklearn.datasets import make_classification
>>> from sklearn import config_context
>>> from sklearn.discriminant_analysis import LinearDiscriminantAnalysis
>>> import cupy.array_api as xp

>>> X_np, y_np = make_classification(random_state=0)
>>> X_cu = xp.asarray(X_np)
>>> y_cu = xp.asarray(y_np)
>>> X_cu.device
<CUDA Device 0>

>>> with config_context(array_api_dispatch=True):
...     lda = LinearDiscriminantAnalysis()
...     X_trans = lda.fit_transform(X_cu, y_cu)
>>> X_trans.device
<CUDA Device 0>
```

After the model is trained, fitted attributes that are arrays will also be from the same Array API namespace as the training data. For example, if CuPy's Array API namespace was used for training, then fitted attributes will be on the GPU. We provide a experimental `_estimator_with_converted_arrays` utility that transfers an estimator attributes from Array API to a ndarray:

```
>>> from sklearn.utils._array_api import _estimator_with_converted_arrays
>>> cupy_to_ndarray = lambda array : array._array.get()
>>> lda_np = _estimator_with_converted_arrays(lda, cupy_to_ndarray)
>>> X_trans = lda_np.transform(X_np)
>>> type(X_trans)
<class 'numpy.ndarray'>
```

11.1.2. Estimators with support for `Array API`-compatible inputs

- `discriminant_analysis.LinearDiscriminantAnalysis` (with `solver="svd"`)

Coverage for more estimators is expected to grow over time. Please follow the dedicated meta-issue on GitHub to track progress.