lathe Documentation

Release 0.2.1

mwilliammyers

Apr 08, 2017

Contents

Basic machine learning tools for BYU CS478.

CHAPTER 1

contents

- genindex
- modindex
- search

lathe

Module Contents

lathe.**bar**(*title, data, xlabel, xticklabels, ylabel, file=None, figsize=None, xlim=None, ylim=None, colors=['b', 'g', 'r', 'c', 'm', 'y', 'k']*)

lathe.**plot**(*title, xdata, ydata, ylabel=None, file=None, figsize=None, xlim=None, ylim=None, colors=['b', 'g', 'r', 'c', 'm', 'y', 'k'], font_size=16*)

lathe.**parse_args**(*parser=<function _parse_args>*)

Parse arguments from the command line.

Parameters **parser** (*function, optional*) – The argument parsing function to use.

Returns The parsed arguments.

Return type argparse.Namespace

See also:

•https://docs.python.org/3.4/library/argparse.html#argparse.Namespace

lathe.**load**(*file_path*, *label_size=0*, *encode_nominal=True*, *one_hot_data=False*, *one_hot_targets=False*, *imputer=None*, *normalizer=None*, *shuffle=False*)

Load an ARFF file.

Parameters

- **file_path** (*str*) – The path of the ARFF formatted file to load.

- **label_size** (*int, optional*) – The number of labels (outputs) the dataset to load has.
- **encode_nominal** (*bool, optional*) – Whether or not to encode nominal atributes as integers.
- **one_hot_data** (*bool, optional*) – Whether or not to use a one-hot encoder for nominal attributes in *data*. Defaults to whatever the value of *encode_nominal* is.
- **one_hot_targets** (*bool, optional*) – Whether or not to use a one-hot encoder for nominal attributes in *targets*.
- **imputer** (*function, optional*) – A 1 arity function that accepts the dataset to impute missing values over. e.g: *sklearn.preprocessing.Imputer().fit_transform*. Defaults to *None*.
- **normalizer** (*function, optional*) – A 1 arity function that accepts the data to be scaled as a parameter and returns the scaled data. e.g: *lathe.minmax_scale*. Defaults to *None*.
- **shuffle** (*bool, optional*) – Whether or not to shuffle the *data*.

Returns Tuple containing (*attributes*, *data*, *targets*). Where *attributes* is a list of tuples containing (attribute_name, attribute_type), *data* are the features and *targets* are the expected output for the dataset.

Return type (list, *numpy.ndarray*, *numpy.ndarray*)

Note: *targets* will be *None* unless *label_size* >= 1.

See also:

- http://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.OneHotEncoder.html
- http://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.Imputer.html
- http://www.cs.waikato.ac.nz/ml/weka/arff.html

lathe.**shuffle**(*features*, *labels*)

lathe.**split**(*features*, *labels*, *percent*)

lathe.**k_fold**(*data*, *n_splits*, *shuffle=False*)

lathe.**minmax_scale**(*data*, *axis=0*)

Transforms features by scaling *data* along *axis* between 0-1.

Parameters

- **data** (*np.ndarray*) – The data to scale.
- **axis** (*int*) – The axis to scale along.

Returns The scaled data.

Return type (*np.ndarray*)

lathe.**measure_error**(*predictions*, *targets*, *rtol=0*, *evaluator=<function mse>*)

lathe.**measure_accuracy**(*predictions*, *targets*, *rtol=0*)

lathe.**evaluate**(*data*, *targets*, *predict_function*, *measure_functions=None*, *rtol=0*, *progress=False*, **args*)

lathe.**get_continuous_index**(*attributes*)

lathe.**get_nominal_index**(*attributes*)

Submodules

lathe.checkpoint

lathe.checkpoint.**load**(*path*)

lathe.checkpoint.**save**(*path*, *checkpoint*)

lathe.metrics

lathe.metrics.**mse**(*predictions*, *targets*)

lathe.metrics.**rmse**(*predictions*, *targets*)

CHAPTER 2

requirements

- python2.7 or python3.3+
- pip (*optional*)

CHAPTER 3

installation

```
pip install lathe
```

CHAPTER 4

usage

```
import lathe

args = lathe.parse_args()
attributes, data, targets = lathe.load(args.file, label_size=1)
```

CHAPTER 5

documentation

http://lathe.readthedocs.io

Python Module Index

l

Index