
xfer

Release 1.1.0

Jun 25, 2019

1	Deep Transfer Learning with Xfer	3
2	Model Handler	13
3	Transfer learning for text categorization	19
4	Xfer with HyperParameter Optimization	25
5	Using ModelHandler with Gluon	35
6	Using Gluon with Xfer	41
7	Repurposing	47
8	Model Handler	69
9	Writing a custom Repurposer	75
10	Indices and tables	81
	Python Module Index	83
	Index	85

Release: 1.1.0

Date: Jun 25, 2019

Xfer is a [Transfer Learning](#) framework written in Python.

Xfer features Repurposers that can be used to take an MXNet model and train a meta-model or modify the model for a new target dataset. To get started with Xfer checkout our introductory tutorial [here](#).

The code can be found on our [Github project page](#). It is open source and provided using the Apache 2.0 license.

Deep Transfer Learning with Xfer

1.1 Transfer learning in 3 lines of code:

```
repurposer = xfer.LrRepurposer(source_model, feature_layer_names=['fc7'])
repurposer.repurpose(train_iterator)
predictions = repurposer.predict_label(test_iterator)
```

Keep reading below to see Xfer in action!

1.2 Overview

1.2.1 What is Xfer?

Xfer is a library that allows quick and easy transfer of knowledge stored in deep neural networks. It can be used for the classification of data of arbitrary numeric format, and can be applied to the common cases of image or text data.

Xfer can be used as a pipeline that spans from extracting features to training a repurposer. The repurposer is then an object that performs classification in the target task.

You can also use individual components of Xfer as part of your own pipeline. For example, you can leverage the feature extractor to extract features from deep neural networks or ModelHandler, which allows for quick building of neural networks, even if you are not an MXNet expert.

1.2.2 How can Xfer help me?

- *Resource efficiency*: you don't have to train big neural networks from scratch.
- *Data efficiency*: by transferring knowledge, you can classify complex data even if you have very few labels.
- *Easy access to neural networks*: you don't need to be an ML ninja in order to leverage the power of neural networks. With Xfer you can easily re-use them or even modify existing architectures and create your own solution.

- *Uncertainty modeling*: With the Bayesian neural network (BNN) or the Gaussian process (GP) repurposers, you can obtain uncertainty in the predictions of the repurposer.
- *Utilities for feature extraction from neural networks*.
- *Rapid prototyping*.

1.2.3 This Demo

In this notebook we demonstrate Xfer in an image classification task. A pre-trained neural network is selected, from which we transfer knowledge for the classification task in the *target* domain. The *target* task is a much smaller set of images that come from a different domain (hand-drawn sketches), therefore the classifier from the *source* task cannot be used as is, without repurposing. Therefore, the aim is to train a new classifier and it is vital to transfer knowledge from the *source* task, due to the extremely scarce *target* dataset. The new classifier for the *target* task is either a meta-model or a modified and fine-tuned clone of the *source* task's neural network.

1.2.4 Components

Xfer is comprised of 2 components:

- `ModelHandler` - Extracts features from pretrained model and performs model manipulation
- `Repurposer` - Repurposes model for target task

1.2.5 Transfer Learning Pipeline

In the following, we demonstrate the Xfer workflow:

1. A data iterator creation
2. A pre-trained model selection (i.e. picking a *source task*)
3. Feature extraction with the `ModelHandler`
4. `Repurposer` used to perform transfer learning from the source task to the target task

First we import or define all relevant modules and utilities.

```
In [1]: import numpy as np
import os
import json
import random
import logging
import glob

import mxnet as mx
from sklearn.model_selection import StratifiedShuffleSplit
from sklearn.metrics import classification_report
from matplotlib import pylab as plt
%matplotlib inline

import xfer

seed=2
random.seed(seed)
np.random.seed(seed)
mx.random.seed(seed)
```

```

logger = logging.getLogger()
logger.setLevel(logging.INFO)

# Change the default option below to test Xfer on other datasets (or use your own!).
TEST_IMAGES = 'test_sketches/' # Options: 'test_images' or 'test_sketches' or 'test_images_s

In [2]: def get_iterators_from_folder(data_dir, train_size=0.6, batchsize=10, label_name='softmax_lab
    """
    Method to create iterators from data stored in a folder with the following structure:
    /data_dir
        /class1
            class1_img1
            class1_img2
            ...
            class1_imgN
        /class2
            class2_img1
            class2_img2
            ...
            class2_imgN
        ...
        /classN
    """
    # assert dir exists
    if not os.path.isdir(data_dir):
        raise ValueError('Directory not found: {}'.format(data_dir))
    # get class names
    classes = [x.split('/')[1] for x in glob.glob(data_dir+'/*')]
    classes.sort()
    fnames = []
    labels = []
    for c in classes:
        # get all the image filenames and labels
        images = glob.glob(data_dir+'/' + c + '/*')
        images.sort()
        fnames += images
        labels += [c]*len(images)
    # create label2id mapping
    id2label = dict(enumerate(set(labels)))
    label2id = dict((v,k) for k, v in id2label.items())

    # get indices of train and test
    sss = StratifiedShuffleSplit(n_splits=2, test_size=None, train_size=train_size, random_state=None)
    train_indices, test_indices = next(sss.split(labels, labels))

    train_img_list = []
    test_img_list = []
    train_labels = []
    test_labels = []
    # create imglist for training and test
    for idx in train_indices:
        train_img_list.append([label2id[labels[idx]], fnames[idx]])
        train_labels.append(label2id[labels[idx]])
    for idx in test_indices:
        test_img_list.append([label2id[labels[idx]], fnames[idx]])
        test_labels.append(label2id[labels[idx]])

    # make iterators
    train_iterator = mx.image.ImageIter(batchsize, (3,224,224), imglist=train_img_list, labels=train_labels)

```

```
        path_root='')
test_iterator = mx.image.ImageIter(batchsize, (3,224,224), imglist=test_img_list, label_
        path_root='')

return train_iterator, test_iterator, train_labels, test_labels, id2label, label2id

def get_images(iterator):
    """
    Returns list of image arrays from iterator
    """
    iterator.reset()
    images = []
    while True:
        try:
            batch = iterator.next().data[0]
            for n in range(batch.shape[0]):
                images.append(batch[n])
        except StopIteration:
            break
    return images

def show_predictions(predictions, images, id2label, uncertainty=None, figsize=(9,1.2), fontsize=12):
    """
    Plots images with predictions as labels. If uncertainty is given then this is plotted below
    series of horizontalbar charts.
    """
    num_rows = 1 if uncertainty is None else 2

    plt.figure(figsize=figsize)
    for cc in range(n):
        plt.subplot(num_rows,n,1+cc)
        plt.tick_params(
            axis='both',          # changes apply to the x-axis
            which='both',        # both major and minor ticks are affected
            bottom=False,        # ticks along the bottom edge are off
            top=False,           # ticks along the top edge are off
            left=False,
            labelleft=False,
            labelbottom=False) # labels along the bottom edge are off
        plt.imshow(np.uint8(images[cc].asnumpy().transpose((1,2,0))))
        plt.title(id2label[predictions[cc]].split(',')[0], fontsize=fontsize)
        plt.axis

    if uncertainty is not None:
        pos = range(len(id2label.values()))
        for cc in range(n):
            plt.subplot(num_rows,n,n+1+cc)
            # Normalize the bars to be 0-1 for better readability.
            xx = uncertainty[cc]
            xx = (xx-min(xx))/(max(xx)-min(xx))
            plt.barh(pos, xx, align='center', height=0.3)
            if cc == 0:
                plt.yticks(pos, id2label.values())
            else:
                plt.gca().set_yticklabels([])
            plt.gca().set_xticklabels([])
            plt.grid(True)
```

1.3 Data Handling

In order for Xfer to process data, it must be given as an MXNet data iterator (`mxnet.io.DataIter`). MXNet expects labels to be sequential integers starting at zero so we have mapped all our string labels to integers to avoid any unexpected behaviours.

The data handling portion of the workflow is made up of the following steps:

- Get iterators
- Get labels
- Get label to idx mapping dictionary

```
In [3]: # We have chosen to split the data into train and test at a 60:40 ratio and use a batchsize of 100
        train_iterator, test_iterator, train_labels, test_labels, id2label, label2id = get_iterators()

INFO:root:Using 1 threads for decoding...
INFO:root:Set environment variable MXNET_CPU_WORKER_NTHREADS to a larger number to use more threads.
INFO:root:ImageIter: loading image list...
INFO:root:Using 1 threads for decoding...
INFO:root:Set environment variable MXNET_CPU_WORKER_NTHREADS to a larger number to use more threads.
INFO:root:ImageIter: loading image list...
```

1.4 Source Model

`ModelHandler` is an Xfer module which handles everything related to the source pre-trained neural network. It can extract features given a target dataset and source model, and it can also manipulate the pre-trained network by adding/removing/freezing layers (we'll see this functionality in the next section). For now, we simply:

- Load MXNet Module from file
- Instantiate `ModelHandler` object with VGG-19 model as source model

The VGG-19 model is a convolutional neural network trained on ImageNet and is good at image classification. Other models trained on ImageNet are likely to be good source models for this task.

```
In [4]: # Download model
        path = 'http://data.mxnet.io/models/imagenet/'
        [mx.test_utils.download(path+'vgg/vgg19-0000.params'),
         mx.test_utils.download(path+'vgg/vgg19-symbol.json')]

INFO:root:vgg19-0000.params exists, skipping download
INFO:root:vgg19-symbol.json exists, skipping download

Out[4]: ['vgg19-0000.params', 'vgg19-symbol.json']

In [5]: source_model = mx.module.Module.load('vgg19', 0, label_names=['prob_label'])
        mh = xfer.model_handler.ModelHandler(source_model)
```

1.5 How well the pre-trained network alone is doing (without repurposing)?

This section will show how well the pre-trained *source* model performs before any repurposing is applied.

```
In [6]: # Get pre-trained model without modifications
        model = mh.get_module(iterator=test_iterator)
```

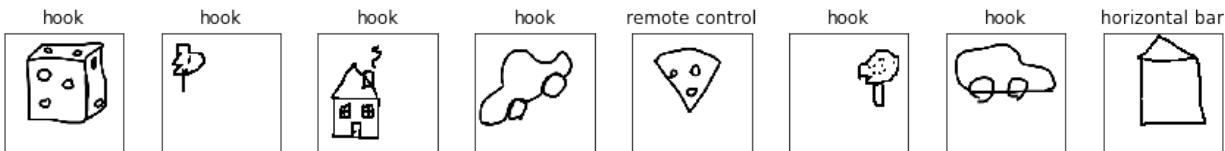
```
# Predict on our test data
predictions = np.argmax(model.predict(test_iterator), axis=1).astype(int)

In [7]: # This utility just allows us to translate image-id's of the imagenet dataset to human-readable
        with open('imagenet1000-class-to-human.json', 'r') as fp:
            imagenet_class_to_human = json.load(fp)

        imagenet_class_to_human = {int(k): v for k, v in imagenet_class_to_human.items()}

In [8]: # Plot all test images along with the predicted labels
        images = get_images(test_iterator)

        show_predictions(predictions, images, imagenet_class_to_human, None, (15, 1.5))
```



The model is performing badly on our sketch images - it thinks most of our drawings are hooks! The reason for this is that the label and image distribution in the target task are different (having come from a different dataset) i.e the model has been trained on photographs of objects and so cannot sensibly classify these sketches. The results would get worse if the source/target dataset mismatch was larger. A **repurposing** step is required to better align the pre-trained model with the target data.

1.6 Repurposing

1.6.1 (a) Repurposing with meta-models

By repurposing with meta models, we use the neural network as a feature extractor and fit a different model on these features.

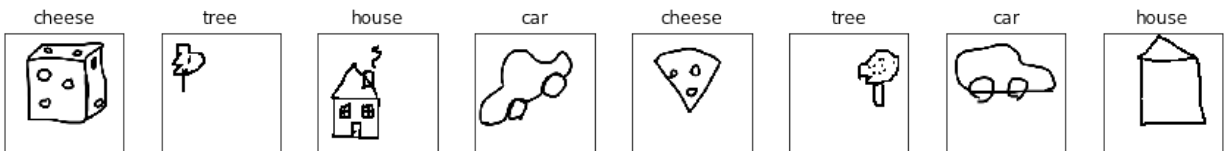
```
In [9]: # Instantiate a Logistic Regression repurposer (other options: SVM; GP; NN, BNN repurposers)
        logging.info("Logistic Regression (LR) Repurposer")
        repLR = xfer.LrRepurposer(source_model=source_model, feature_layer_names=['fc7'])
        repLR.repurpose(train_iterator)
        predictionsLR = repLR.predict_label(test_iterator)

        logging.info("LR Repurposer - Classification Results")
        print(classification_report(test_labels, predictionsLR, target_names=list(id2label.values())))

INFO:root:Logistic Regression (LR) Repurposer
INFO:root:Extracting features from layers: fc7
INFO:root:Processed batch 1
INFO:root:Processed batch 2
INFO:root:Processed batch 3
INFO:root:Processed batch 4
INFO:root:Processed batch 5
INFO:root:Processed batch 6
/anaconda/envs/matplotlib-backend/lib/python3.5/site-packages/sklearn/linear_model/sag.py:326: Conv
INFO:root:Extracting features from layers: fc7
INFO:root:Processed batch 1
INFO:root:Processed batch 2
INFO:root:Processed batch 3
INFO:root:Processed batch 4
INFO:root:LR Repurposer - Classification Results
```

	precision	recall	f1-score	support
tree	1.000	0.750	0.857	4
car	1.000	1.000	1.000	4
cheese	1.000	1.000	1.000	4
house	0.800	1.000	0.889	4
avg / total	0.950	0.938	0.937	16

```
In [10]: show_predictions(predictionsLR, images, id2label, None, (15,1.5))
```



1.6.2 (b) Fine-tuning Neural Network repurposer

Neural network repurposers will:

- Modify the pretrained neural network architecture by adding and removing layers
- Retrain the network with certain layers held fixed or randomised

```
In [11]: # Choose which layers of the model to fix during training - more fixed layers lead to faster
fixed_layers = ['conv1_1', 'conv1_2', 'conv2_1', 'conv2_2', 'conv3_1', 'conv3_2', 'conv3_3', 'conv3_4',
               'conv4_1', 'conv4_2', 'conv4_3', 'conv4_4', 'conv5_1', 'conv5_2', 'conv5_3', 'conv5_4']
# Choose which layers of the model to randomise before training - we may want to forget some
# this model knows
random_layers = []
```

```
repNN = xfer.NeuralNetworkRandomFreezeRepurposer(source_model, target_class_count=4, fixed_layers=fixed_layers)
repNN.repurpose(train_iterator)
predictionsNN = repNN.predict_label(test_iterator)
logging.info("NN Repurposer - Classification Results")
print(classification_report(test_labels, predictionsNN, target_names=list(id2label.values())))
```

```
INFO:root:fc8, prob deleted from model top
INFO:root:Added new_fully_connected_layer, prob to model top
WARNING:root:Already bound, ignoring bind()
/anaconda/envs/matplotlib-backend/lib/python3.5/site-packages/mxnet/module/base_module.py:488: UserWarning:
INFO:root:Epoch[0] Train-accuracy=0.541667
INFO:root:Epoch[0] Time cost=33.846
INFO:root:Epoch[1] Train-accuracy=1.000000
INFO:root:Epoch[1] Time cost=22.451
INFO:root:Epoch[2] Train-accuracy=1.000000
INFO:root:Epoch[2] Time cost=24.884
INFO:root:Epoch[3] Train-accuracy=1.000000
INFO:root:Epoch[3] Time cost=23.454
INFO:root:Epoch[4] Train-accuracy=1.000000
INFO:root:Epoch[4] Time cost=24.850
INFO:root:NN Repurposer - Classification Results
```

	precision	recall	f1-score	support
tree	1.000	0.750	0.857	4
car	1.000	1.000	1.000	4

cheese	1.000	1.000	1.000	4
house	0.800	1.000	0.889	4
avg / total	0.950	0.938	0.937	16

The neural network repurposer will likely not be great if the target dataset is extremely small.

1.6.3 (c) Repurposing with probability and uncertainty

Two repurposers offer well-calibrated probability for predictions: GPRepurposer and BNNRepurposer. Here we explore the former (the latter can be used for non-tiny datasets).

```
In [12]: # Instantiate a GP repurposer
         repGP = xfer.GPRepurposer(source_model, feature_layer_names=['fc6'], apply_l2_norm=True)
         repGP.repurpose(train_iterator)

         logging.info("GP Repurposer - Classification Results")
         uncertaintyGP = repGP.predict_probability(test_iterator)
         predictionsGP = np.argmax(uncertaintyGP, axis=1)

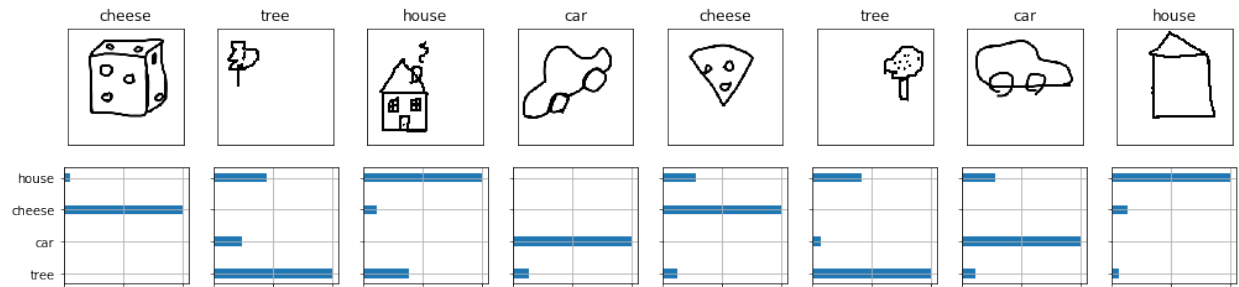
         print(classification_report(test_labels, predictionsGP,
                                     target_names=list(id2label.values()), digits=3))

INFO:root:Extracting features from layers: fc6
INFO:root:Processed batch 1
INFO:root:Processed batch 2
INFO:root:Processed batch 3
INFO:root:Processed batch 4
INFO:root:Processed batch 5
INFO:root:Processed batch 6
INFO:GP:initializing Y
INFO:GP:initializing inference method
INFO:GP:adding kernel and likelihood as parameters
INFO:GP:initializing Y
INFO:GP:initializing inference method
INFO:GP:adding kernel and likelihood as parameters
INFO:GP:initializing Y
INFO:GP:initializing inference method
INFO:GP:adding kernel and likelihood as parameters
INFO:GP:initializing Y
INFO:GP:initializing inference method
INFO:GP:adding kernel and likelihood as parameters
INFO:root:GP Repurposer - Classification Results
INFO:root:Extracting features from layers: fc6
INFO:root:Processed batch 1
INFO:root:Processed batch 2
INFO:root:Processed batch 3
INFO:root:Processed batch 4
```

	precision	recall	f1-score	support
tree	1.000	0.750	0.857	4
car	1.000	1.000	1.000	4
cheese	1.000	1.000	1.000	4
house	0.800	1.000	0.889	4
avg / total	0.950	0.938	0.937	16

The code below will plot the predictions and the probability for each class.

```
In [13]: show_predictions(predictionsGP, images, id2label, uncertaintyGP, (17,3.8))
```



We not only have predictions from this model but we can also see the uncertainty in the model for any given prediction which allows us to make better decisions on our data.

1.6.4 Other repurposers

We have seen the use of LrRepurposer, NeuralNetworkRandomFreezeRepurposer and GpRepurposer. Other repurposers offered are: SvmRepurposer, BnnRepurposer, NeuralNetworkFineTuneRepurposer.

You can also *write your own repurposer*

1.7 Using Xfer on your own data

All you need to do is generate your own data iterator and use it instead of the iterators used above.

For more details see the [API documentation](#)

```
In [ ]:
```


ModelHandler is a utility class for manipulating and inspecting MXNet models. It can be used to:

- Add and remove layers from an existing model or “freeze” selected layers
- Discover information such as layer names and types
- Extract features from pretrained models

In this tutorial, we will demonstrate some of the key capabilities of ModelHandler.

2.1 Initialisation

```
In [1]: import mxnet as mx
import logging

import xfer

logger = logging.getLogger()
logger.setLevel(logging.INFO)

In [2]: # Download vgg19 (trained on imagenet)
path = 'http://data.mxnet.io/models/imagenet/'
[mx.test_utils.download(path+'vgg/vgg19-0000.params'),
mx.test_utils.download(path+'vgg/vgg19-symbol.json')]

INFO:root:vgg19-0000.params exists, skipping download
INFO:root:vgg19-symbol.json exists, skipping download

Out[2]: ['vgg19-0000.params', 'vgg19-symbol.json']

In [3]: source_model = mx.module.Module.load('vgg19', 0, label_names=['prob_label'])

# The ModelHandler constructor takes an MXNet Module as input
mh = xfer.model_handler.ModelHandler(source_model)
```

2.2 Model Inspection

2.2.1 Layer names

```
In [4]: print(mh.layer_names)
```

```
['conv1_1', 'relu1_1', 'conv1_2', 'relu1_2', 'pool1', 'conv2_1', 'relu2_1', 'conv2_2', 'relu2_2', 'p
```

2.2.2 Layer Types

Given the name of a layer, this function returns the type.

```
In [5]: print(mh.get_layer_type('relu5_2'))
        print(mh.get_layer_type('flatten_0'))
        print(mh.get_layer_type('fc7'))
        print(mh.get_layer_type('conv5_3'))
        print(mh.get_layer_type('prob'))
```

Activation

Flatten

FullyConnected

Convolution

SoftmaxOutput

ModelHandler can be used to get a list of layers that are of a specific type.

```
In [6]: import xfer.model_handler.consts as consts
```

```
        print(mh.get_layer_names_matching_type('Convolution'))
        print(mh.get_layer_names_matching_type('Pooling'))
        print(mh.get_layer_names_matching_type('Activation'))
        print(mh.get_layer_names_matching_type('BatchNorm'))
```

```
['conv1_1', 'conv1_2', 'conv2_1', 'conv2_2', 'conv3_1', 'conv3_2', 'conv3_3', 'conv3_4', 'conv4_1',
['pool1', 'pool2', 'pool3', 'pool4', 'pool5']
['relu1_1', 'relu1_2', 'relu2_1', 'relu2_2', 'relu3_1', 'relu3_2', 'relu3_3', 'relu3_4', 'relu4_1',
[]
```

2.2.3 Architecture Visualization

```
In [7]: mh.visualize_net()
```

2.3 Feature Extraction

ModelHandler makes it easy to extract features from a dataset using a pretrained model.

By passing an MXNet DataIterator and a list of the layers to extract features from the `get_layer_output()` method will return a feature dictionary and an ordered list of labels.

```
In [8]: imglist = [[0, 'test_images/accordion/accordion_1.jpg'], [0, 'test_images/accordion/accordion_4.jpg'], [0, 'test_images/accordion/accordion_5.jpg'],
                  [1, 'test_images/ant/ant_2.jpg'], [1, 'test_images/ant/ant_3.jpg'], [1, 'test_images/ant/ant_4.jpg'],
                  [2, 'test_images/anchor/anchor_1.jpg'], [2, 'test_images/anchor/anchor_2.jpg'], [2, 'test_images/anchor/anchor_3.jpg'],
                  [2, 'test_images/anchor/anchor_4.jpg'], [2, 'test_images/anchor/anchor_5.jpg'], [2, 'test_images/anchor/anchor_6.jpg']]
```

```

        [3, 'test_images/airplanes/airplanes_2.jpg'], [3, 'test_images/airplanes/airplanes_5.jpg']]
    iterator = mx.img.ImageIter(imglist=imglist, batch_size=4, path_root='', data_shape=(3, 224, 224))

INFO:root:Using 1 threads for decoding...
INFO:root:Set enviroment variable MXNET_CPU_WORKER_NTHREADS to a larger number to use more threads.
INFO:root:ImageIter: loading image list...

In [9]: features, labels = mh.get_layer_output(data_iterator=iterator, layer_names=['fc6', 'fc8'])

        print('Shape of output from fc6:', features['fc6'].shape)
        print('Shape of output from fc8:', features['fc8'].shape)
        print('Labels:', labels)

        print('Subset of example feature:', features['fc8'][0,:100])

INFO:root:Extracting features from layers: fc6 fc8
INFO:root:Processed batch 1
INFO:root:Processed batch 2
INFO:root:Processed batch 3
INFO:root:Processed batch 4
INFO:root:Processed batch 5

Shape of output from fc6: (20, 4096)
Shape of output from fc8: (20, 1000)
Labels: [0 0 0 0 0 1 1 1 1 1 2 2 2 2 2 3 3 3 3 3]
Subset of example feature: [-2.5173218 -3.0606608 -0.18566455 -1.195584 -1.6561157 -0.3466432
-0.27523482 -3.0833778 -3.987122 -4.711858 -2.5934136 -1.1821984
 0.06423883 -3.5403426 0.36486915 -2.0515091 -3.3651357 0.47566187
-0.93592185 -0.53005326 -2.7707744 -1.2674817 -2.3202353 -0.33125317
-1.5847255 -4.2490544 -4.2170153 -5.6999183 -2.6653297 -3.4800928
-4.693992 -3.4104934 -3.673527 -4.2224913 -0.29074478 -6.513745
-4.4927287 -4.5361094 -2.549627 2.1703975 -1.3125131 -2.1347325
-3.761081 -2.3712082 -3.8052034 -1.6259451 -1.68117 -1.481512
-2.2081814 -1.5731778 -1.287838 1.2327844 -3.9466934 -3.9385183
-0.87836707 -2.9489741 -3.4411037 -4.030957 -1.4967936 -3.7117271
-2.2397022 -3.325867 -2.8145652 -0.63274264 -3.2671835 -1.8046627
-3.0445974 -1.5151932 -2.7235372 6.5550556 -1.62281 -1.5104069
 1.3592944 0.8891826 -0.14048216 -1.0063077 -1.5578198 -0.45763612
-2.0689113 3.2839453 -2.0749338 -4.179339 -0.49392343 -0.5244163
-1.9723302 -0.07367857 -2.2878125 -0.96980214 -2.8648748 -2.6847577
-3.5610118 -3.7286394 -4.5710897 -4.949738 -0.80546796 -5.8007493
-3.260846 -6.434879 -4.7502995 -4.953493 ]

```

These features can be used for training a meta-model or for clustering as shown in this example.

```

In [10]: from sklearn.decomposition import PCA
        from sklearn.cluster import KMeans
        import matplotlib.pyplot as plt
        import numpy as np
        %matplotlib inline

        reduced_data = PCA(n_components=2).fit_transform(features['fc8'])

        kmeans = KMeans(init='k-means++', n_clusters=4, n_init=10)
        kmeans.fit(reduced_data)

        h=0.1

        x_min, x_max = reduced_data[:, 0].min() - 1, reduced_data[:, 0].max() + 1
        y_min, y_max = reduced_data[:, 1].min() - 1, reduced_data[:, 1].max() + 1

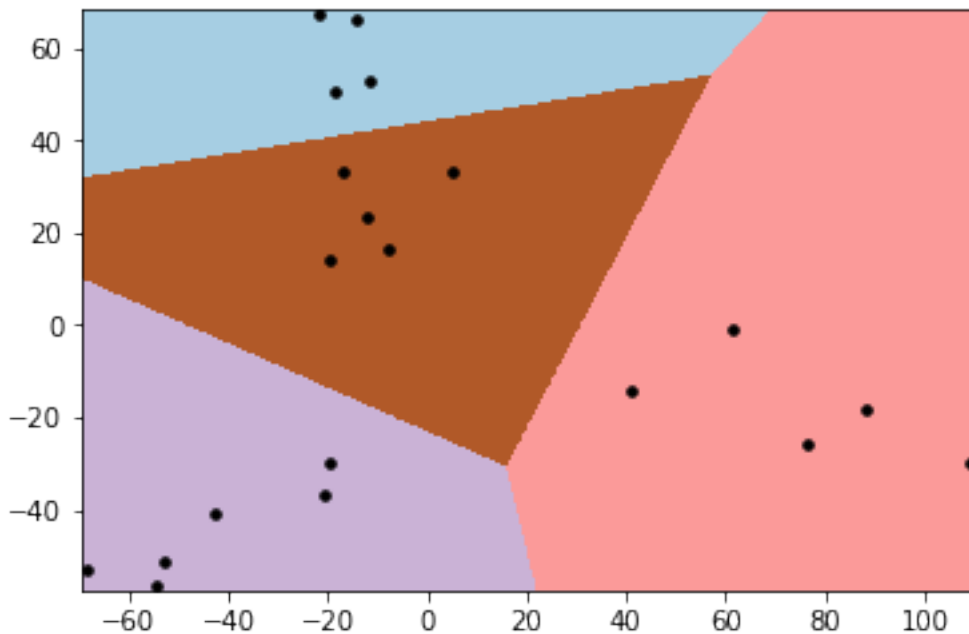
```

```
xx, yy = np.meshgrid(np.arange(x_min, x_max, h), np.arange(y_min, y_max, h))

Z = kmeans.predict(np.c_[xx.ravel(), yy.ravel()])
Z = Z.reshape(xx.shape)

plt.imshow(Z, interpolation='nearest',
           extent=(xx.min(), xx.max(), yy.min(), yy.max()),
           cmap=plt.cm.Paired,
           aspect='auto', origin='lower')
plt.plot(reduced_data[:, 0], reduced_data[:, 1], 'k.', markersize=7)
```

Out [10]: [



2.4 Model Manipulation

Modifying models in MXNet can be problematic because symbols are held as graphs. This means that modifying the input of the model requires the graph to be reconstructed above any changes made. ModelHandler takes care of this for you which means that adding and removing layers from either end of a neural network can be done with 1-2 lines of code.

2.4.1 Remove layers

```
In [11]: # Dropping 4 layers from the top of the layer hierarchy (where top = output)
         mh.drop_layer_top(4)
         print(mh.layer_names)
```

INFO:root:relu7, drop7, fc8, prob deleted from model top

```
['conv1_1', 'relu1_1', 'conv1_2', 'relu1_2', 'pool1', 'conv2_1', 'relu2_1', 'conv2_2', 'relu2_2', 'p
```

```
In [12]: # Dropping a layer from the bottom of the layer hierarchy (where bottom = input)
         mh.drop_layer_bottom(1)
         print(mh.layer_names)
```

```
INFO:root:conv1_1 deleted from model bottom
```

```
['relu1_1', 'conv1_2', 'relu1_2', 'pool1', 'conv2_1', 'relu2_1', 'conv2_2', 'relu2_2', 'pool2', 'conv
```

2.4.2 Add layers

Layers can be added to models by first defining the layer with an `mxnet.symbol` object and using `add_layer_top()` or `add_layer_bottom()` to add the layer to the model.

```
In [13]: # define layer symbols
fc = mx.sym.FullyConnected(name='fullyconnected1', num_hidden=4)
softmax = mx.sym.SoftmaxOutput(name='softmax')
conv1 = mx.sym.Convolution(name='convolution1', kernel=(20,20), num_filter=64)

# Add layer to the bottom of the layer hierarchy (where bottom = input)
mh.add_layer_bottom([conv1])
# Add layer to the top of the layer hierarchy (where top = output)
mh.add_layer_top([fc, softmax])

print(mh.layer_names)
```

```
INFO:root:Added convolution1 to model bottom
```

```
INFO:root:Added fullyconntected1, softmax to model top
```

```
['convolution1', 'relu1_1', 'conv1_2', 'relu1_2', 'pool1', 'conv2_1', 'relu2_1', 'conv2_2', 'relu2_2']
```

Once a model has been modified, `ModelHandler` can be used to return an `MXNet Module` which can then be used for training.

There is an option to specify parameters which should stay fixed during training or should be randomised before training to allow different modes of transfer learning.

```
In [14]: # In this case, the conv1_1 layer will stay fixed during training and the layers fc6 and fc7
         mod = mh.get_module(iterator,
                               fixed_layer_parameters=mh.get_layer_parameters(['conv1_1']),
                               random_layer_parameters=mh.get_layer_parameters(['fc6', 'fc7']))
```

```
In [15]: iterator.reset()
          mod.fit(iterator, num_epoch=5)
```

```
WARNING:root:Already bound, ignoring bind()
```

```
/anaconda/envs/xfer_env/lib/python3.5/site-packages/mxnet/module/base_module.py:488: UserWarning: Parameter allow_missing=allow_missing, force_init=force_init)
```

```
INFO:root:Epoch[0] Train-accuracy=0.100000
```

```
INFO:root:Epoch[0] Time cost=58.709
```

```
INFO:root:Epoch[1] Train-accuracy=0.250000
```

```
INFO:root:Epoch[1] Time cost=56.932
```

```
INFO:root:Epoch[2] Train-accuracy=0.250000
```

```
INFO:root:Epoch[2] Time cost=52.735
```

```
INFO:root:Epoch[3] Train-accuracy=0.250000
```

```
INFO:root:Epoch[3] Time cost=61.472
```

```
INFO:root:Epoch[4] Train-accuracy=0.250000
```

```
INFO:root:Epoch[4] Time cost=58.052
```

We now have a trained model ready to be used for prediction. This new model isn't very useful but demonstrates the concept - to train a better model, use more data and experiment with combinations of fixed and random layers.

Now you have seen what `ModelHandler` can do, you should try it out for yourself!

For more details see the [API docs](#)

In []:

Transfer learning for text categorization

In this notebook we showcase how to use Xfer to tackle a simple task of transfer learning for text categorization. To that end, we use the 20 newsgroups text dataset (<http://qwone.com/~jason/20Newsgroups/>) which is comprised of a collection of 20K newsgroups posts.

We use a Convolutional Neural Network (CNN) pre-trained on a subset of 13 classes (~12K instances). For the target task, we assume that we have access to a much smaller dataset with 100 posts from the remaining 7 categories. This is a common situation in many real world applications where the number of categories grows with time as we collect new data. In this case, we will always start with a low number of labelled instances for the new categories (this is the cold-start problem).

In this scenario, training a NN from scratch on the target task is not feasible: due to the scarcity of labeled instances (100 in this case) and the large number of parameters, the model will be prone to overfitting. Instead, we will use Xfer to transfer the knowledge from the source model and propose a data efficient classifier.

```
In [1]: import warnings
        warnings.filterwarnings('ignore')

        import re
        import numpy as np
        import pickle
        from collections import Counter
        import itertools
        import numpy as np
        import mxnet as mx

        import xfer

        mx.random.seed(1)
        np.random.seed(1)

        %matplotlib inline
        import matplotlib.pyplot as plt

In [2]: config = {
        "source_vocab": "NewsGroupsSourceVocabulary.pickle",
        "model_prefix_source": 'NewsGroupsSourceModel',
```

```
"num_epoch_source": 100,  
"batch_size": 100,  
"num_train": 100,  
"context": mx.cpu(),  
}
```

3.1 A) Load the pre-trained model

Before building the target dataset, we load the pre-trained model into a `mxnet Module`. In this case, the pre-trained model is a CNN that was trained with instances belonging to the following 13 categories that are not used in the target task:

- `comp.graphics`
- `comp.os.ms-windows.misc`
- `comp.sys.ibm.pc.hardware`
- `comp.windows.x`
- `misc.forsale`
- `rec.motorcycles`
- `rec.sport.baseball`
- `sci.crypt`
- `sci.med`
- `sci.space`
- `talk.politics.mideast`
- `talk.politics.misc`
- `talk.religion.misc`

```
In [3]: import zipfile  
        with zipfile.ZipFile("{}-{:04}.params.zip".format(config["model_prefix_source"], config["num_train_source"]), "r") as zip_ref:  
            zip_ref.extractall()  
  
In [4]: sym, arg_params, aux_params = mx.model.load_checkpoint(config["model_prefix_source"], config["num_train_source"])  
        mx.viz.plot_network(sym)
```

3.2 B) Load the target dataset into an iterator

We define a helper function to download the dataset. The 7 classes used for the target task are the following:

- `alt.atheism`
- `comp.sys.mac.hardware`
- `rec.autos`
- `rec.sport.hockey`
- `sci.electronics`
- `soc.religion.christian`

- talk.politics.guns

```
In [5]: def download_dataset():
    from sklearn.datasets import fetch_20newsgroups
    categories=['alt.atheism',
               'comp.sys.mac.hardware',
               'rec.autos',
               'rec.sport.hockey',
               'sci.electronics',
               'soc.religion.christian',
               'talk.politics.guns'
    ]

    newsgroups_train = fetch_20newsgroups(subset='train', categories=categories)
    newsgroups_test = fetch_20newsgroups(subset='test', categories=categories)

    x_text = np.concatenate((newsgroups_train.data, newsgroups_test.data), axis=0)
    labels = np.concatenate((newsgroups_train.target, newsgroups_test.target))

    return x_text, labels, categories
```

In addition, we use two helper classes to create the corpus: * **Vocabulary**: It creates the lexicon for a given corpus. In addition, it provides a basic string cleaning function based on regular expressions. * **Corpus**: Given a corpus (text and labels) and a Vocabulary object, it converts the text instances into a numerical format. In particular, it uses the provided vocabulary object to tokenize and clean the text instances. Then, it pads the sentences using `max_length/fix_length` and the padding symbol defined in the vocabulary object. Finally, each token is encoded into a one-hot vector using the vocabulary. In addition, it provides a helper function to build the training and test sets.

```
In [6]: class Vocabulary(object):
    def __init__(self, sentences, padding_word="</s>", unknown_word="</ukw>"):
        self.padding_word = padding_word
        self.unknown_word = unknown_word
        sentences = [self.clean_str(sent).split(" ") for sent in sentences]
        self.max_length = max(len(x) for x in sentences)
        self.word_counts = Counter(itertools.chain(*sentences))
        self.id2word = [x[0] for x in self.word_counts.most_common()]
        self.id2word.append(self.padding_word)
        self.id2word.append(self.unknown_word)
        self.word2id = {x: i for i, x in enumerate(self.id2word)}

        print('Vocabulary size', len(self.id2word))

    def clean_str(self, string):
        string = re.sub(r"^[A-Za-z0-9(),;!?\\']", " ", string)
        contractions = ["'t", "'ve", "'d", "'s", "'ll", "'m", "'er"]
        punctuations = [",", ";", "!", "?", "\\", "("]
        for ee in contractions + punctuations:
            string = re.sub(r"{}".format(ee), " {}".format(ee), string)
        return string.strip().lower()

In [7]: class Corpus(object):
    def __init__(self, sentences, labels, vocabulary, max_length=None, fix_length=None):
        self.vocabulary = vocabulary
        self.max_length = max_length
        self.fix_length = fix_length
        sentences = [self.vocabulary.clean_str(sent).split(" ") for sent in sentences]
        sentences_padded = self._pad_sentences(sentences, self.vocabulary.padding_word, self
        x = []
        for sentence in sentences_padded:
            x.append([self.vocabulary.word2id.get(word, self.vocabulary.word2id[self.vocabulary
```

```
self.x = np.array(x)
self.y = np.array(labels)

print('Data shape:', self.x.shape)
print('Vocabulary size', len(vocabulary.id2word))
print('Maximum number words per sentence', self.x.shape[1])
print('Number of labels', len(np.unique(self.y)))

def _pad_sentences(self, sentences, padding_word="</s>", max_length=None, fix_length=None):
    sequence_length = max(len(x) for x in sentences)
    if max_length is not None:
        sequence_length = min(sequence_length, max_length)
    if fix_length is not None:
        sequence_length = fix_length
    padded_sentences = []
    for i in range(len(sentences)):
        sentence = sentences[i]
        if len(sentence) > sequence_length:
            sentence = sentence[0:sequence_length]
        num_padding = sequence_length - len(sentence)
        new_sentence = sentence + [padding_word] * num_padding
        padded_sentences.append(new_sentence)

    return padded_sentences

def split_train_test(self, num_train):
    shuffle_indices = np.random.permutation(np.arange(len(self.y)))
    x_shuffled = self.x[shuffle_indices]
    y_shuffled = self.y[shuffle_indices]
    x_train, x_dev = x_shuffled[:num_train], x_shuffled[num_train:]
    y_train, y_dev = y_shuffled[:num_train], y_shuffled[num_train:]

    print('Train/Test split: %d/%d' % (len(y_train), len(y_dev)))
    print('Train shape:', x_train.shape)
    print('Test shape:', x_dev.shape)

    return x_train, x_dev, y_train, y_dev
```

We import the vocabulary object of the source model

```
In [8]: with open(config["source_vocab"], 'rb') as handle:
        vocab_source = pickle.load(handle)
```

We build the dataset for the target task and split it in training/test sets. Notice that we use only 100 instances for the training set since we care about scenarios in which the target dataset is small compared to the source dataset.

```
In [9]: x_text, labels, categories = download_dataset()
        corpus = Corpus(x_text, labels, vocab_source, fix_length=arg_params["vocab_embed_weight"].shape[1])
        x_train, x_dev, y_train, y_dev = corpus.split_train_test(config["num_train"])
```

```
Data shape: (6642, 300)
Vocabulary size 149371
Maximum number words per sentence 300
Number of labels 7
Train/Test split: 100/6542
Train shape: (100, 300)
Test shape: (6542, 300)
```

We finally load the training/test datasets into mxnet iterators:

```
In [10]: train_iter = mx.io.NDArrayIter(x_train, y_train, config['batch_size'], shuffle=True)
        val_iter = mx.io.NDArrayIter(x_dev, y_dev, config['batch_size'], shuffle=False)
```

3.3 C) How well we predict without repurposing?

We create a new mxnet Module, bind the training/test iterators and load the parameters of the pre-trained model.

```
In [11]: mod = mx.mod.Module(symbol=sym,
                             context=config["context"],
                             data_names=['data'],
                             label_names=['softmax_label'])
mod.bind(data_shapes=train_iter.provide_data, label_shapes=train_iter.provide_label)
mod.set_params(arg_params, aux_params)

In [12]: mod.save_checkpoint("NewsGroupsSourceModelCompressed",1)
```

If we use the pre-trained network directly, we obtain a poor performance since it was trained for a different task (different set of classes).

```
In [13]: y_predicted = mod.predict(val_iter).asnumpy().argmax(axis=1)
print("Accuracy score before repurposing: {}".format(np.mean((y_predicted == y_dev))))

Accuracy score before repurposing: 0.07169061449098135
```

3.4 D) Repurposing

For this demo, we use one class of repurposers called `MetaModelRepurposer`. These are repurposers that use the pre-trained model as a feature extractor and fit a different model using the extracted features. In particular, we use an `LrRepurposer` which fits a logistic regression using the extracted features.

```
In [14]: lr_repurposer = xfer.LrRepurposer(mod, ["dropout0"])
lr_repurposer.repurpose(train_iter)
y_predicted = lr_repurposer.predict_label(val_iter)

In [15]: print("Accuracy score after repurposing: {}".format(np.mean((y_predicted == y_dev))))

Accuracy score after repurposing: 0.5889636196881688
```

Training the model from scratch results in an accuracy of approximately 0.5 and it is much slower to train having to resort to GPUs to speed up the process.

Finally, we show the full pipeline to categorize a new test example. In this case, we have used the first paragraphs of the wikipedia article for “**Automotive Electronics**”. (Feel free to replace it with your own example!!!)

```
In [16]: test_example = ["""
Automotive electronics are electronic systems used in vehicles, including engine management,

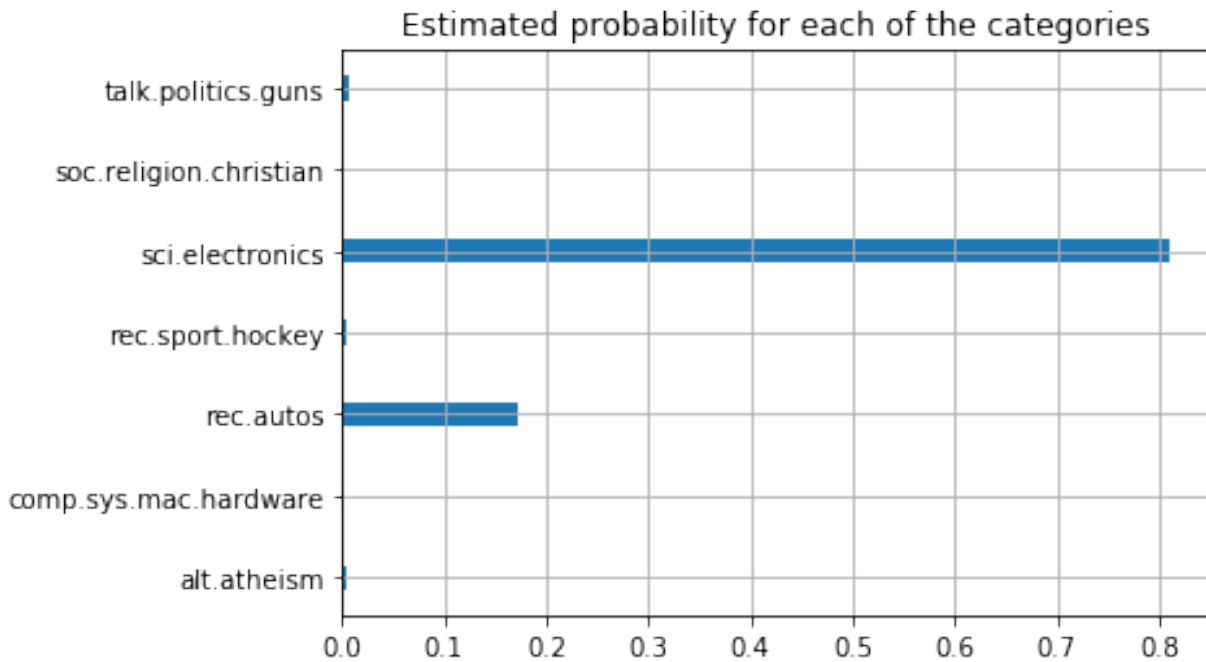
Electronic systems have become an increasingly large component of the cost of an automobile,

The earliest electronics systems available as factory installations were vacuum tube car radios.
"""]

In [17]: corpus = Corpus(test_example, [1], vocab_source, fix_length=arg_params["vocab_embed_weight"])
Data shape: (1, 300)
Vocabulary size 149371
Maximum number words per sentence 300
Number of labels 1

In [18]: test_example_iter = mx.io.NDArrayIter(corpus.x, corpus.y, len(corpus.y), shuffle=False)
In [19]: y_predicted_prob = lr_repurposer.predict_probability(test_example_iter)
In [20]: plt.barh(list(range(y_predicted_prob.shape[1])), y_predicted_prob[0,:], align='center', height=0.5)
plt.yticks(list(range(y_predicted_prob.shape[1])), categories)
```

```
plt.title("Estimated probability for each of the categories")  
plt.grid(True)
```



We can see that the main two identified topics are electronics and autos, which is what we would expect given the title of the article.

In []:

Xfer with HyperParameter Optimization

When training neural networks, hyperparameters may have to be tuned to improve accuracy metrics. The purpose of this notebook is to demonstrate how to do *HyperParameter Optimization* (HPO) when repurposing neural networks in [Xfer](#). Here, we use [emukit](#) to do HPO through Bayesian Optimization.

Note that depending on number of epochs, the target data set and transferability between source and target tasks, the default hyperparameter settings in Xfer could give desired results and HPO may not be required. If someone wants to try HPO, this notebook shows how to do it using emukit.

```
In [1]: import warnings
        warnings.filterwarnings("ignore")

        import logging
        logging.disable(logging.WARNING)

        import gc
        import glob
        import os
        import random
        import time

        import emukit
        import mxnet as mx
        import numpy as np
        from sklearn.model_selection import StratifiedShuffleSplit
        import xfer

        from matplotlib import pylab as plt
        %matplotlib inline
```

4.1 Utility methods

```
In [2]: def set_random_seeds():
        seed = 1234
```

```
np.random.seed(seed)
mx.random.seed(seed)
random.seed(seed)

def get_iterators(data_dir, train_size=0.3, validation_size=0.3, test_size=0.4, batch_size=1,
                  label_name='softmax_label', data_name='data', random_state=1):
    """
    Method to create iterators from data stored in a folder with the following structure:
    /data_dir
      /class1
        class1_img1 ... class1_imgN
      /class2
        class2_img1 ... class2_imgN
      ...
      /classN
    """
    set_random_seeds()
    # Assert data_dir exists
    if not os.path.isdir(data_dir):
        raise ValueError('Directory not found: {}'.format(data_dir))
    # Get class names
    classes = [x.split('/')[0] for x in glob.glob(data_dir+'/*')]
    classes.sort()
    fnames = []
    labels = []
    for c in classes:
        # Get all the image filenames and labels
        images = glob.glob(data_dir+'/' + c + '/*')
        images.sort()
        fnames += images
        labels += [c]*len(images)
    # Create label2id mapping
    id2label = dict(enumerate(set(labels)))
    label2id = dict((v,k) for k, v in id2label.items())

    # Split training(train+validation) and test data
    sss = StratifiedShuffleSplit(n_splits=2, test_size=None, train_size=train_size+validation_size)
    train_indices, test_indices = next(sss.split(labels, labels))

    # Training data (train+validation)
    train_validation_images = []
    train_validation_labels = []
    for idx in train_indices:
        train_validation_images.append([label2id[labels[idx]], fnames[idx]])
        train_validation_labels.append(label2id[labels[idx]])

    # Test data
    test_images = []
    test_labels = []
    for idx in test_indices:
        test_images.append([label2id[labels[idx]], fnames[idx]])
        test_labels.append(label2id[labels[idx]])

    # Separate validation set and train set
    train_percent = train_size / (train_size+validation_size)
    sss_1 = StratifiedShuffleSplit(n_splits=2, test_size=None, train_size=train_percent, random_state=random_state)
    train_indices, validation_indices = next(sss_1.split(train_validation_labels, train_validation_images))
    train_images = []
    train_labels = []
```

```

for idx in train_indices:
    train_images.append(train_validation_images[idx])
    train_labels.append(train_validation_labels[idx])
validation_images = []
validation_labels = []
for idx in validation_indices:
    validation_images.append(train_validation_images[idx])
    validation_labels.append(train_validation_labels[idx])

# Create iterators
train_iterator = mx.image.ImageIter(batch_size, (3,224,224), imglist=train_images, label_
                                data_name=data_name, path_root='')
validation_iterator = mx.image.ImageIter(batch_size, (3,224,224), imglist=validation_images,
                                data_name=data_name, path_root='')
train_validation_iterator = mx.image.ImageIter(batch_size, (3,224,224), imglist=train_validation_images,
                                label_name=label_name, data_name=data_name, path_root='')
test_iterator = mx.image.ImageIter(batch_size, (3,224,224), imglist=test_images, label_name=label_name,
                                data_name=data_name, path_root='')

return train_iterator, validation_iterator, train_validation_iterator, test_iterator, id2label

def get_labels(iterator):
    """ Return labels from data iterator """
    iterator.reset()
    labels = []
    while True:
        try:
            labels = labels + iterator.next().label[0].asnumpy().astype(int).tolist()
        except StopIteration:
            break
    return labels

def get_images(iterator):
    """ Return list of image arrays from iterator """
    iterator.reset()
    images = []
    while True:
        try:
            batch = iterator.next().data[0]
            for n in range(batch.shape[0]):
                images.append(batch[n])
        except StopIteration:
            break
    return images

def show_predictions(predictions, images, id2label, figsize=(15,1.5), fontsize=12, n=None):
    """ Display images along with predicted labels """
    n = len(images) if n is None else n
    num_rows = 1
    plt.figure(figsize=figsize)
    for cc in range(n):
        plt.subplot(num_rows,n,1+cc)
        plt.tick_params(
            axis='both',          # changes apply to the x-axis
            which='both',        # both major and minor ticks are affected
            bottom=False,        # ticks along the bottom edge are off
            top=False,           # ticks along the top edge are off
            left=False,
            labelleft=False,

```

```
        labelbottom=False) # labels along the bottom edge are off
plt.imshow(np.uint8(images[cc].asnumpy().transpose((1,2,0))))
plt.title(id2label[predictions[cc]].split(',')[0], fontsize=fontsize)
plt.axis
```

4.1.1 A) Source model

In Transfer Learning, the model from which knowledge is transferred is called the source model. Here, we use `vgg19` model from [MXNet Model Zoo](#) as the source model. `vgg19` is a convolutional neural network trained on the ImageNet dataset which contains 1 million natural images categorized into 1000 classes.

```
In [3]: # Download source model
path = 'http://data.mxnet.io/models/imagenet/'
[mx.test_utils.download(path+'vgg/vgg19-0000.params'), mx.test_utils.download(path+'vgg/vgg19-0000.params')]

# Load source model from file
source_model = mx.module.Module.load('vgg19', 0, label_names=['prob_label'])
```

4.1.2 B) Target data

Target data is a much smaller dataset with 40 images categorized into 4 classes from a different domain (hand-drawn sketches). We'll demonstrate how to use Xfer along with HPO to learn classifying this target data by transferring knowledge from `vgg19` model.

```
In [4]: TARGET_DATA_DIR = 'test_sketches'
set_random_seeds()
train_iterator, validation_iterator, train_validation_iterator, test_iterator, id2label = get_iter_and_labels(TARGET_DATA_DIR)
train_labels = get_labels(train_iterator)
validation_labels = get_labels(validation_iterator)
test_labels = get_labels(test_iterator)
train_validation_labels = get_labels(train_validation_iterator)

print('Number of train images: {}'.format(len(train_labels)))
print('Number of validation images: {}'.format(len(validation_labels)))
print('Number of test images: {}'.format(len(test_labels)))
```

```
Number of train images: 12
Number of validation images: 12
Number of test images: 16
```

4.2 How are these data sets used?

During HPO, we train the model using the training data and evaluate the hyperparameters with the validation data. Once we find an optimized learning rate, we do a final train of the model using both the training data and the validation data, and report precision on our withheld testing data.

4.2.1 C) Repurpose without HPO

This section demonstrates how to repurpose the source model to target data with default hyperparameters.

```
In [5]: # Default optimizer, learning rate and number of epochs used in Xfer to train neural network
DEFAULT_OPTIMIZER = 'sgd'
DEFAULT_LEARNING_RATE = 0.01
DEFAULT_OPTIMIZER_PARAMS = {'learning_rate': DEFAULT_LEARNING_RATE}
```

```

DEFAULT_NUM_EPOCHS = 4

TARGET_CLASS_COUNT = 4 # 4 classes of sketch images are used for this demo (car, cheese, house, tree)
CONTEXT_FUNCTION = mx.cpu # 'mx.gpu' or 'mx.cpu' (MXNet context function to train neural network)

# Layers to freeze or randomly initialize during neural network training.
# For this demo, we freeze the first 16 convolutional layers transferred from 'vgg19' model.
FIXED_LAYERS = ['conv1_1', 'conv1_2', 'conv2_1', 'conv2_2', 'conv3_1', 'conv3_2', 'conv3_3', 'conv3_4',
                'conv4_1', 'conv4_2', 'conv4_3', 'conv4_4', 'conv5_1', 'conv5_2', 'conv5_3', 'conv5_4']
RANDOM_LAYERS = []

# Method to repurpose neural network with given hyperparameters
def train_and_predict(source_model, train_data_iterator, test_data_iterator, optimizer, optimizer_params):
    set_random_seeds()
    repurposer = xfer.NeuralNetworkRandomFreezeRepurposer(source_model,
                                                           optimizer=optimizer,
                                                           optimizer_params=optimizer_params,
                                                           target_class_count=TARGET_CLASS_COUNT,
                                                           fixed_layers=FIXED_LAYERS,
                                                           random_layers=RANDOM_LAYERS,
                                                           context_function=CONTEXT_FUNCTION,
                                                           num_epochs=DEFAULT_NUM_EPOCHS)

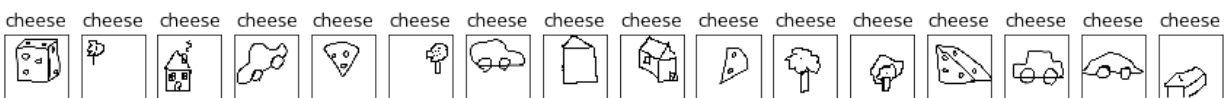
    repurposer.repurpose(train_data_iterator)
    predictions = repurposer.predict_label(test_data_iterator)
    return predictions

# Train neural network with default hyperparameters
predictions = train_and_predict(source_model, train_validation_iterator, test_iterator,
                                DEFAULT_OPTIMIZER, DEFAULT_OPTIMIZER_PARAMS)
precision_default = np.mean(predictions == test_labels)
print('Trained neural network with default hyperparameters. Precision: {}'.format(precision_default))

# Display test images and predictions
test_images = get_images(test_iterator)
show_predictions(predictions, test_images, id2label)

```

Trained neural network with default hyperparameters. Precision: 0.25



Note that the quality of the results may vary due to randomness e.g. in the initialization of the neural network weights. However, the idea is that, since the precision varies with different choices of the learning rate and in certain cases the default learning rate might not be the best. We therefore wish to find a good learning rate automatically rather than with trial-and-error. HPO helps us achieve this, as demonstrated below.

4.2.2 D) Repurpose with HPO: Optimizing learning rate

4.3 i) Declare the hyperparameter to optimize and its domain

```

In [6]: # Learning rate is the hyperparameter we will optimize here
        # We allow emukit to operate in a normalized domain [0,1] and map the value to a desired log
        # This helps emukit to learn a smooth underlying function in fewer iterations
        from emukit.core import ParameterSpace, ContinuousParameter

```

```
emukit_param_space = ParameterSpace([ContinuousParameter('learning_rate', 0, 1)])

# Method to map a value given by emukit in [0, 1] to a desired range of learning rate
def map_learning_rate(source_value):
    if(source_value < 0 or source_value > 1):
        raise ValueError('source_value must be in the range [0,1]')

    # We explore learning rate in the range [1e-6 , 1e-1]. You can choose a different range to
    # Log scale is used here because it is an intuitive way to explore learning rates
    # For example, if 1e-2 doesn't work, we tend to explore 1e-3 or 1e-1 which is a jump in 1
    log_learning_rate_start = -6 # 1e-6 in linear scale
    log_learning_rate_end = -1 # 1e-1 in linear scale

    log_span = abs(log_learning_rate_end - log_learning_rate_start)
    log_mapped_value = log_learning_rate_start + (source_value * log_span)
    mapped_value = 10 ** log_mapped_value # Convert from log scale to linear scale
    return mapped_value
```

4.4 ii) Define an objective function to optimize the hyperparameter

```
In [7]: def get_hyperparameters_from_config(config):
    """
    Extract hyperparameters from input configuration provided by emukit.
    Refer the caller 'hpo_objective_function' for more details.
    """
    learning_rate = map_learning_rate(config[0]) # Map learning_rate value given by emukit to
    optimizer = DEFAULT_OPTIMIZER # Using default optimizer here i.e. 'sgd'
    return optimizer, learning_rate

def hpo_objective_function(config_matrix):
    """
    Objective function to optimize the hyperparameters for
    This method is called by emukit during the optimization loop
    to get outputs of objective function for different input configurations

    We train a neural network with given hyperparameters and return (1-precision) on validation.
    You can choose to optimize for a different measure and create the objective function accordingly.
    Here, we consider one hyperparameter (learning_rate) to optimize precision

    Note: config_matrix has m rows and n columns
    m denotes the number of experiments to run i.e. each row would contain input configuration
    n denotes the number of hyperparameters (e.g. 2 columns for learning_rate and batch_size)
    """
    # Output of objective function for each input configuration
    function_output = np.zeros((config_matrix.shape[0], 1))

    # For each input configuration, train a neural network and calculate accuracy on validation
    for idx, config in enumerate(config_matrix):
        optimizer, learning_rate = get_hyperparameters_from_config(config)

        # Train neural network with the mapped learning rate and get predictions on validation
        predictions = train_and_predict(source_model=source_model,
                                       train_data_iterator=train_iterator,
                                       test_data_iterator=validation_iterator,
                                       optimizer = optimizer,
                                       optimizer_params = {'learning_rate': learning_rate})
```

```

        # Calculate precision on validation set and update function_output with (1-precision)
        precision = np.mean(predictions == validation_labels)
        function_output[idx][0] = (1.0 - precision) # (1-precision) to keep a minimization
        print('learning_rate: {}'.format(learning_rate), optimizer, precision)
        gc.collect()

    return function_output

```

4.5 iii) Define a model using GPy for emukit to optimize

```

In [8]: # Notice that our initial learning rate value corresponds to 0.8 on [0, 1] scale
        map_learning_rate(0.8) == DEFAULT_LEARNING_RATE

```

```

Out[8]: True

```

```

In [9]: set_random_seeds()

import GPy
X = np.array([[0.8]])
Y = np.array([[1.0 - precision_default]])

gpy_model = GPy.models.GPRegression(X, Y)

```

4.6 iv) Initialize a Bayesian optimizer using emukit

```

In [10]: from emukit.bayesian_optimization.loops import BayesianOptimizationLoop
         from emukit.model_wrappers import GPyModelWrapper

         emukit_model = GPyModelWrapper(gpy_model)
         hyperparameter_optimizer = BayesianOptimizationLoop(emukit_param_space, emukit_model)

```

4.7 v) Run optimization loop to identify a better learning rate for our objective function

```

In [11]: NUM_ITERATIONS_TO_RUN = 7

         hyperparameter_optimizer.run_loop(hpo_objective_function, NUM_ITERATIONS_TO_RUN)
         results = hyperparameter_optimizer.get_results()

Optimization restart 1/1, f = 1.1312564608151783
learning_rate: 1e-06. optimizer: sgd. precision: 0.5
Optimization restart 1/1, f = 0.9815790760721594
learning_rate: 1e-06. optimizer: sgd. precision: 0.5
Optimization restart 1/1, f = -6.963249113616735
learning_rate: 1.604696232278503e-06. optimizer: sgd. precision: 0.5
Optimization restart 1/1, f = -9.048731845639432
learning_rate: 1.2681487716415876e-06. optimizer: sgd. precision: 0.5
Optimization restart 1/1, f = -16.18851622078782
learning_rate: 1.260324119615607e-06. optimizer: sgd. precision: 0.5
Optimization restart 1/1, f = -24.10206487648798
learning_rate: 6.481443855100692e-06. optimizer: sgd. precision: 0.5833333333333334
Optimization restart 1/1, f = -20.776248579222724

```

```
learning_rate: 0.00011946457911772839. optimizer: sgd. precision: 1.0
Optimization restart 1/1, f = -21.117596675753305
```

4.8 Optimized learning rate

```
In [12]: # Take hyperparameters that minimized the objective function output
x_best = results.minimum_location
optimized_learning_rate = map_learning_rate(x_best[0])
precision = 1.0 - results.minimum_value # Objective was to minimize 1-precision
print('Optimized learning rate: {}'.format(optimized_learning_rate))
```

```
Optimized learning rate: 0.00011946457911772839. Precision on validation data: 1.0
```

```
In [13]: gpy_model
```

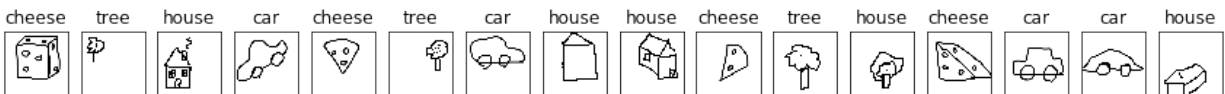
```
Out[13]: <GPy.models.gp_regression.GPRegression at 0x146c27470>
```

Note that the optimized learning rate found is from the iterations run so far. Based on time available, one can run more iterations which may help in obtaining a more optimized learning rate.

4.9 Precision on test data with optimal learning rate

```
In [14]: # Train neural network with optimal learning rate and get predictions on test data
# Along with training data, validation data is also used to train the final model
predictions = train_and_predict(source_model=source_model,
                                train_data_iterator=train_validation_iterator, # train on
                                test_data_iterator=test_iterator, # predict on test data
                                optimizer = DEFAULT_OPTIMIZER,
                                optimizer_params = {'learning_rate': optimized_learning_rate})
precision_optimized = np.mean(predictions == test_labels)
print('Optimized learning rate: {}'.format(optimized_learning_rate))
show_predictions(predictions, test_images, id2label)
```

```
Optimized learning rate: 0.00011946457911772839. Precision on test data: 0.9375
```



4.9.1 E) Repurpose with HPO: Optimizing multiple hyperparameters

The following section can be used for reference when someone wants to optimize multiple hyperparameters to repurpose models using Xfer. Here, the hyperparameters chosen are: 1. Optimizer for neural network (sgd or adam). 2. Learning rate.

Note that running more iterations could be useful here because there are more combination of values to explore.

```
# Choose the hyperparameters and specify the domain
from emukit.core import CategoricalParameter, OneHotEncoding
p1 = ContinuousParameter('learning_rate', 0, 1)
p2 = CategoricalParameter('optimizer', OneHotEncoding(['sgd', 'adam']))
space_with_two_params = ParameterSpace([p1, p2])

# Override this method to extract the optimizer in addition to learning_rate from_
↳ emukit config
```

(continues on next page)

(continued from previous page)

```

def get_hyperparameters_from_config(config):
    """
    Extract hyperparameters from input configuration provided by emukit.
    Refer the caller 'hpo_objective_function' for more details.
    """
    learning_rate = map_learning_rate(config[0]) # Map learning_rate value given by_
    ↪emukit to the desire range
    optimizer = p2.encoding.get_category(config[1:]) # Using optimizer given by_
    ↪emukit
    return optimizer, learning_rate

# Initialize emukit with new domain, create the model and run optimization
set_random_seeds()

X = np.array([[0.8] + p2.encoding.get_encoding(DEFAULT_OPTIMIZER)])
Y = np.array([[1.0 - precision_default]])
gpy_model = GPy.models.GPRegression(X, Y)

emukit_model = GPyModelWrapper(gpy_model)
hyperparameter_optimizer2 = BayesianOptimizationLoop(space_with_two_params, emukit_
    ↪model)

hyperparameter_optimizer2.run_loop(hpo_objective_function, NUM_ITERATIONS_TO_RUN)
results2 = hyperparameter_optimizer2.get_results()

# Take hyperparameters that minimized the objective function output
x_best2 = results2.minimum_location
best_learning_rate = map_learning_rate(x_best2[0])
best_optimizer = p2.encoding.get_category(x_best2[1:])
precision2 = 1.0 - results2.minimum_value # Objective was to minimize 1-precision
print('Optimized learning rate: {}. Optimizer: {}. Precision on validation data: {}'.
    .format(best_learning_rate, best_optimizer, precision2))

# Train neural network with optimal (learning rate, optimizer) and get predictions on_
    ↪test data
# Along with training data, validation data is also used to train the final model
predictions2 = train_and_predict(source_model=source_model,
    train_data_iterator=train_validation_iterator, #_
    ↪train on (train + validation) set
    test_data_iterator=test_iterator, # predict on_
    ↪test data
    optimizer = best_optimizer,
    optimizer_params = {'learning_rate': best_learning_
    ↪rate}))
precision_optimized2 = np.mean(predictions2 == test_labels)
print('Precision on test data after optimization: ' + str(precision_optimized2))
show_predictions(predictions2, test_images, id2label)

```

In []:

Using ModelHandler with Gluon

MXNet's Gluon framework allows Neural Networks to be written under an imperative paradigm. ModelHandler is currently based around the symbolic graph implementation of MXNet and as a result, models written in Gluon cannot directly be used.

If the model is written in Gluon using [HybridBlocks](#) (i.e. if the network consists entirely of predefined MXNet layers) then the model can be compiled as a symbolic graph using the command `.hybridize()`.

The Gluon defined model can then be converted to a symbol and set of parameters which can then be loaded as an MXNet Module and used with ModelHandler.

In this demo, we will show that you can define a model in Gluon using code from the [Gluon MNIST demo](#) and then convert it to a Module and use ModelHandler.

```
In [1]: import mxnet as mx
        from mxnet import gluon
        from mxnet.gluon import nn
        from mxnet import autograd as ag

        import os

        # Fixing the random seed
        mx.random.seed(42)
```

5.1 Train model in Gluon

5.1.1 Define model in Gluon

```
In [2]: mnist = mx.test_utils.get_mnist()

In [3]: batch_size = 100
        train_data = mx.io.NDArrayIter(mnist['train_data'], mnist['train_label'], batch_size, shuffle=True)
        val_data = mx.io.NDArrayIter(mnist['test_data'], mnist['test_label'], batch_size)
```

```
In [4]: # define network
net = nn.HybridSequential()
with net.name_scope():
    net.add(nn.Dense(128, activation='relu'))
    net.add(nn.Dense(64, activation='relu'))
    net.add(nn.Dense(10))

net.hybridize()

In [5]: gpus = mx.test_utils.list_gpus()
ctx = [mx.gpu()] if gpus else [mx.cpu(0), mx.cpu(1)]
net.initialize(mx.init.Xavier(magnitude=2.24), ctx=ctx)
trainer = gluon.Trainer(net.collect_params(), 'sgd', {'learning_rate': 0.02})
```

5.1.2 Training

```
In [6]: %%time
epoch = 10
# Use Accuracy as the evaluation metric.
metric = mx.metric.Accuracy()
softmax_cross_entropy_loss = gluon.loss.SoftmaxCrossEntropyLoss()
for i in range(epoch):
    # Reset the train data iterator.
    train_data.reset()
    # Loop over the train data iterator.
    for batch in train_data:
        # Splits train data into multiple slices along batch_axis
        # and copy each slice into a context.
        data = gluon.utils.split_and_load(batch.data[0], ctx_list=ctx, batch_axis=0)
        # Splits train labels into multiple slices along batch_axis
        # and copy each slice into a context.
        label = gluon.utils.split_and_load(batch.label[0], ctx_list=ctx, batch_axis=0)
        outputs = []
        # Inside training scope
        with ag.record():
            for x, y in zip(data, label):
                z = net(x)
                # Computes softmax cross entropy loss.
                loss = softmax_cross_entropy_loss(z, y)
                # Backpropagate the error for one iteration.
                loss.backward()
                outputs.append(z)
        # Updates internal evaluation
        metric.update(label, outputs)
        # Make one step of parameter update. Trainer needs to know the
        # batch size of data to normalize the gradient by 1/batch_size.
        trainer.step(batch.data[0].shape[0])
    # Gets the evaluation result.
    name, acc = metric.get()
    # Reset evaluation result to initial state.
    metric.reset()
    print('training acc at epoch {}: {}'.format(i, name, acc))

training acc at epoch 0: accuracy=0.7816
training acc at epoch 1: accuracy=0.89915
training acc at epoch 2: accuracy=0.9134666666666666
training acc at epoch 3: accuracy=0.9225833333333333
training acc at epoch 4: accuracy=0.9305666666666667
```

```

training acc at epoch 5: accuracy=0.9366666666666666
training acc at epoch 6: accuracy=0.9418166666666666
training acc at epoch 7: accuracy=0.94585
training acc at epoch 8: accuracy=0.9495333333333333
training acc at epoch 9: accuracy=0.9532333333333334
CPU times: user 43.1 s, sys: 4.18 s, total: 47.3 s
Wall time: 29.7 s

```

5.2 Testing

```

In [7]: # Use Accuracy as the evaluation metric.
metric = mx.metric.Accuracy()
# Reset the validation data iterator.
val_data.reset()
# Loop over the validation data iterator.
for batch in val_data:
    # Splits validation data into multiple slices along batch_axis
    # and copy each slice into a context.
    data = gluon.utils.split_and_load(batch.data[0], ctx_list=ctx, batch_axis=0)
    # Splits validation label into multiple slices along batch_axis
    # and copy each slice into a context.
    label = gluon.utils.split_and_load(batch.label[0], ctx_list=ctx, batch_axis=0)
    outputs = []
    for x in data:
        outputs.append(net(x))
    # Updates internal evaluation
    metric.update(label, outputs)
print('validation acc: {}={}'.format(*metric.get()))
assert metric.get()[1] > 0.94

validation acc: accuracy=0.9527

```

5.3 Convert Gluon model to Module

Adapted from snippet found *here* <<https://github.com/apache/incubator-mxnet/issues/9374>> ‘__

From the Gluon model, the symbol and parameters are extracted and used to define an Module object.

```

In [8]: def block2symbol(block):
    data = mx.sym.Variable('data')
    sym = block(data)
    args = {}
    auxs = {}
    for k, v in block.collect_params().items():
        args[k] = mx.nd.array(v.data().asnumpy())
        auxs[k] = mx.nd.array(v.data().asnumpy())
    return sym, args, auxs

In [9]: def symbol2mod(sym, args, auxs, data_iter):
    mx_sym = mx.sym.SoftmaxOutput(data=sym, name='softmax')
    model = mx.mod.Module(symbol=mx_sym, context=mx.cpu(),
        label_names=['softmax_label'])
    model.bind( data_shapes = data_iter.provide_data,
        label_shapes = data_iter.provide_label )
    model.set_params(args, auxs)
    return model

```

```
In [10]: sym_params = block2symbol(net)
In [11]: mod = symbol2mod(*sym_params, train_data)
```

5.3.1 Alternative Method

Serialise Gluon model to file using `.export()`.

Load the serialised model as an MXNet Module with `Module.load()` so that xfer can be used.

```
In [12]: # model_name = 'gluon-model'
# net.export(model_name)

# mod = mx.mod.Module.load(model_name, 0, label_names=[])
# os.remove(model_name+'-symbol.json')
# os.remove(model_name+'-0000.params')
```

5.4 Apply ModelHandler

Now we can load the model into ModelHandler and use it to visualise the model, return the layer names, extract features and much more!

```
In [13]: import xfer
In [14]: mh = xfer.model_handler.ModelHandler(mod)
In [15]: # Show architecture of model
mh.visualize_net()

In [16]: mh.layer_names
Out[16]: ['hybridsequential0_dense0_fwd',
          'hybridsequential0_dense0_relu_fwd',
          'hybridsequential0_dense1_fwd',
          'hybridsequential0_dense1_relu_fwd',
          'hybridsequential0_dense2_fwd',
          'softmax']

In [17]: # Get output from intermediate layers of the model
mh.get_layer_output(train_data, ['hybridsequential0_dense1_fwd'])
Out[17]: (OrderedDict([('hybridsequential0_dense1_fwd',
                        array([[ 1.93497527e+00,  2.40295935e+00,  1.16074115e-01, ...,
                                -4.74348217e-02, -3.76087427e-03,  1.39985621e+00],
                                [ 2.15391922e+00,  1.97971451e+00,  4.61517543e-01, ...,
                                2.28680030e-01, -8.29489648e-01,  9.69915807e-01],
                                [ 2.06626105e+00,  4.06703472e+00,  7.65578270e-01, ...,
                                3.74726385e-01,  1.03201318e+00, -5.41208267e-01],
                                ...,
                                [ 2.55671740e+00,  4.17255354e+00,  5.60081601e-01, ...,
                                5.68660349e-02, -1.58825326e+00,  1.59997427e+00],
                                [ 2.30686831e+00,  2.34434009e+00, -5.84015131e-01, ...,
                                3.16424906e-01, -1.08476102e-01,  6.86561584e-01],
                                [ 9.71719801e-01,  1.08340001e+00,  1.72682357e+00, ...,
                                -2.98302293e-01,  1.48507738e+00, -7.40276098e-01]]], dtype=float32)),
          array([8, 8, 6, ..., 8, 8, 4]))

In [18]: mh.get_layer_type('hybridsequential0_dense0_relu_fwd')
```

```
Out[18]: 'Activation'
```

```
In [19]: # Add/Remove layers from model output
mh.drop_layer_top(2)
mh.add_layer_top([mx.sym.FullyConnected(num_hidden=30),
                  mx.sym.Activation(act_type='relu'),
                  mx.sym.FullyConnected(num_hidden=10),
                  mx.sym.SoftmaxOutput())])
mh.visualize_net()

In [20]: # Add/remove layers from model input
mh.add_layer_bottom([mx.sym.Convolution(kernel=(2,2), num_filter=10)])
mh.visualize_net()
```

```
In [ ]:
```

Using Gluon with Xfer

This notebook demonstrates how to use neural networks defined and trained with Gluon as source models for Transfer Learning with Xfer.

TL;DR Gluon models can be used with Xfer provided they use [HybridBlocks](#) so that the symbol can be extracted.

This demo is a dummy example where a CNN source model is trained on [MNIST](#) using Gluon and then repurposed for MNIST again. This is obviously redundant but shows the steps required to use Gluon with Xfer.

```
In [1]: import numpy as np
import mxnet as mx
from mxnet import nd, autograd, gluon
mx.random.seed(1)

import time
from sklearn.metrics import classification_report
from scipy import io as scipyio
import urllib.request
import zipfile
import os
import logging

import xfer
```

6.1 Train CNN with gluon

Using code taken from [The Straight Dope](#)

```
In [2]: ctx = mx.cpu()

In [3]: batch_size = 64
num_inputs = 784
num_outputs = 10
def transform(data, label):
    return nd.transpose(data.astype(np.float32), (2,0,1))/255, label.astype(np.float32)
```

```
train_data = gluon.data.DataLoader(gluon.data.vision.MNIST(train=True, transform=transform),
                                   batch_size, shuffle=True)
test_data = gluon.data.DataLoader(gluon.data.vision.MNIST(train=False, transform=transform),
                                  batch_size, shuffle=False)

In [4]: num_fc = 512
net = gluon.nn.HybridSequential()
with net.name_scope():
    net.add(gluon.nn.Conv2D(channels=20, kernel_size=5, activation='relu'))
    net.add(gluon.nn.MaxPool2D(pool_size=2, strides=2))
    net.add(gluon.nn.Conv2D(channels=50, kernel_size=5, activation='relu'))
    net.add(gluon.nn.MaxPool2D(pool_size=2, strides=2))
    # The Flatten layer collapses all axis, except the first one, into one axis.
    net.add(gluon.nn.Flatten())
    net.add(gluon.nn.Dense(num_fc, activation="relu"))
    net.add(gluon.nn.Dense(num_outputs))

In [5]: net.collect_params().initialize(mx.init.Xavier(magnitude=2.24), ctx=ctx)

In [6]: softmax_cross_entropy = gluon.loss.SoftmaxCrossEntropyLoss()

In [7]: trainer = gluon.Trainer(net.collect_params(), 'sgd', {'learning_rate': .1})

In [8]: net.hybridize()

In [9]: def evaluate_accuracy(data_iterator, net):
    acc = mx.metric.Accuracy()
    for i, (data, label) in enumerate(data_iterator):
        data = data.as_in_context(ctx)
        label = label.as_in_context(ctx)
        output = net(data)
        predictions = nd.argmax(output, axis=1)
        acc.update(preds=predictions, labels=label)
    return acc.get()[1]

In [10]: epochs = 1
smoothing_constant = .01

for e in range(epochs):
    start_time_train = time.time()
    for i, (data, label) in enumerate(train_data):
        data = data.as_in_context(ctx)
        label = label.as_in_context(ctx)
        with autograd.record():
            output = net(data)
            loss = softmax_cross_entropy(output, label)
        loss.backward()
        trainer.step(data.shape[0])

        #####
        # Keep a moving average of the losses
        #####
        curr_loss = nd.mean(loss).asscalar()
        moving_loss = (curr_loss if ((i == 0) and (e == 0))
                       else (1 - smoothing_constant) * moving_loss + smoothing_constant * curr_loss)
    end_time_train = time.time()

    start_time_eval = time.time()
    test_accuracy = evaluate_accuracy(test_data, net)
    train_accuracy = evaluate_accuracy(train_data, net)
    end_time_eval = time.time()
```

```
epoch_time = end_time_train - start_time_train
eval_time = end_time_eval - start_time_eval
print("Epoch {}. \nLoss: {}, Train_acc {}, Test_acc {}, Epoch_time {}, Eval_time {}".format
```

Epoch 0.

Loss: 0.11107716270094219, Train_acc 0.9745833333333334, Test_acc 0.9742, Epoch_time 54.263786792755

6.2 Load MNIST dataset

Load MNIST into data iterators

```
In [19]: mnist = mx.test_utils.get_mnist()
```

```
In [20]: train_iter = mx.io.NDArrayIter(mnist['train_data'], mnist['train_label'], batch_size, shuffle=True)
        val_iter = mx.io.NDArrayIter(mnist['test_data'], mnist['test_label'], batch_size)
```

6.3 Convert Gluon model net to Module

Adapted from snippet found 'here' <<https://github.com/apache/incubator-mxnet/issues/9374>> '___

From the Gluon model, the symbol and parameters are extracted and used to define an Module object.

```
In [21]: def block2symbol(block):
        data = mx.sym.Variable('data')
        sym = block(data)
        args = {}
        auxs = {}
        for k, v in block.collect_params().items():
            args[k] = mx.nd.array(v.data().asnumpy())
            auxs[k] = mx.nd.array(v.data().asnumpy())
        return sym, args, auxs

In [22]: def symbol2mod(sym, args, auxs, data_iter):
        mx_sym = mx.sym.SoftmaxOutput(data=sym, name='softmax')
        model = mx.mod.Module(symbol=mx_sym, context=mx.cpu(),
                               label_names=['softmax_label'])
        model.bind( data_shapes = data_iter.provide_data,
                    label_shapes = data_iter.provide_label )
        model.set_params(args, auxs)
        return model

In [23]: sym_params = block2symbol(net)

In [24]: net_mod = symbol2mod(*sym_params, train_iter)
```

6.3.1 Alternative Method

Serialise Gluon model to file using `.export()`.

Load the serialised model as an MXNet Module with `Module.load()` so that xfer can be used.

```
In [25]: # model_name = 'gluon-model'
        # net.export(model_name)

        # mod = mx.mod.Module.load(model_name, 0, label_names=[])
        # os.remove(model_name+'-symbol.json')
        # os.remove(model_name+'-0000.params')
```

6.3.2 Inspect Module

```
In [26]: mh = xfer.model_handler.ModelHandler(net_mod)
```

```
In [27]: mh.layer_names
```

```
Out[27]: ['hybridsequential0_conv0_fwd',
          'hybridsequential0_conv0_relu_fwd',
          'hybridsequential0_pool0_fwd',
          'hybridsequential0_conv1_fwd',
          'hybridsequential0_conv1_relu_fwd',
          'hybridsequential0_pool1_fwd',
          'hybridsequential0_flatten0_reshape0',
          'hybridsequential0_dense0_fwd',
          'hybridsequential0_dense0_relu_fwd',
          'hybridsequential0_dense1_fwd',
          'softmax']
```

6.4 Neural Network Repurposer

```
In [28]: repFT = xfer.NeuralNetworkFineTuneRepurposer(source_model=net_mod,
                                                         transfer_layer_name='hybridsequential0_dense0',
                                                         target_class_count=26, num_epochs=2)
```

```
In [29]: repFT.repurpose(train_iter)
```

```
WARNING:root:Already bound, ignoring bind()
/anaconda/envs/xfer-env/lib/python3.6/site-packages/mxnet/module/base_module.py:488: UserWarning: Parameter
allow_missing=allow_missing, force_init=force_init)
```

```
In [30]: predictionsFT = repFT.predict_label(val_iter)
```

```
In [32]: print(classification_report(mnist['test_label'], predictionsFT,
                                     digits=3))
```

	precision	recall	f1-score	support
0	0.960	0.990	0.975	980
1	0.984	0.989	0.986	1135
2	0.968	0.965	0.967	1032
3	0.967	0.972	0.970	1010
4	0.977	0.971	0.974	982
5	0.975	0.975	0.975	892
6	0.974	0.966	0.970	958
7	0.970	0.959	0.964	1028
8	0.966	0.966	0.966	974
9	0.966	0.954	0.960	1009
avg / total	0.971	0.971	0.971	10000

6.5 Meta-model Repurposer

```
In [33]: repLR = xfer.LrRepurposer(source_model=net_mod, feature_layer_names=['hybridsequential0_dense0'])
```

```
In [34]: repLR.repurpose(train_iter)
```

```
/anaconda/envs/xfer-env/lib/python3.6/site-packages/sklearn/linear_model/sag.py:326: ConvergenceWarning:
  "the coef_ did not converge", ConvergenceWarning)
```

```
In [35]: predictionsLR = repLR.predict_label(val_iter)
```

```
In [36]: print(classification_report(mnist['test_label'], predictionsLR,  
                                     digits=3))
```

	precision	recall	f1-score	support
0	0.990	0.993	0.991	980
1	0.991	0.996	0.993	1135
2	0.985	0.989	0.987	1032
3	0.987	0.989	0.988	1010
4	0.992	0.990	0.991	982
5	0.979	0.982	0.980	892
6	0.990	0.984	0.987	958
7	0.983	0.985	0.984	1028
8	0.987	0.986	0.986	974
9	0.989	0.977	0.983	1009
avg / total	0.987	0.987	0.987	10000

```
In [ ]:
```


Base Classes:

<code>xfer.Repurposer</code>	Base Class for repurposers that train models using Transfer Learning (source_model -> target_model).
<code>xfer.MetaModelRepurposer</code>	Base class for repurposers that extract features from layers in source neural network (Transfer) and train a meta-model using the extracted features (Learn).
<code>xfer.NeuralNetworkRepurposer</code>	Base class for repurposers that create a target neural network from a source neural network through Transfer Learning.

7.1 xfer.Repurposer

class `xfer.Repurposer` (source_model: `mxnet.module.module.Module`, context_function=<function `cpu`>, num_devices=1)

Bases: `object`

Base Class for repurposers that train models using Transfer Learning (source_model -> target_model).

Parameters

- **source_model** (`mxnet.mod.Module`) – Source neural network to do transfer learning from.
- **context_function** (function(int)->:class:`mx.context.Context`) – MXNet context function that provides device type context.
- **num_devices** (`int`) – Number of devices to use with context_function.

Methods

<code>__init__</code>	Initialize self.
<code>deserialize</code>	Abstract method.
<code>get_params</code>	Get parameters of repurposer that are in the constructor's argument list.
<code>predict_label</code>	Abstract method.
<code>predict_probability</code>	Abstract method.
<code>repurpose</code>	Abstract method.
<code>save_repurposer</code>	Serialize the repurposed model (source_model, target_model and supporting info) and save it to given file_path.
<code>serialize</code>	Abstract method.

repurpose (*train_iterator: mxnet.io.io.DataIter*)
Abstract method.

predict_probability (*test_iterator: mxnet.io.io.DataIter*)
Abstract method.

predict_label (*test_iterator: mxnet.io.io.DataIter*)
Abstract method.

get_params ()
Get parameters of repurposer that are in the constructor's argument list.

Return type `dict`

serialize (*file_prefix*)
Abstract method.

deserialize (*input_dict*)
Abstract method.

save_repurposer (*model_name, model_directory=""*, *save_source_model=None*)
Serialize the repurposed model (source_model, target_model and supporting info) and save it to given file_path.

Parameters

- **model_name** (*str*) – Name to save repurposer to.
- **model_directory** (*str*) – File directory to save repurposer in.
- **save_source_model** (*boolean*) – Flag to choose whether to save repurposer source model. Will use default if set to None. (MetaModelRepurposer default: True, NeuralNetworkRepurposer default: False)

7.2 xfer.MetaModelRepurposer

```
class xfer.MetaModelRepurposer (source_model:          mxnet.module.module.Module,    fea-  
                                ture_layer_names,      context_function=<function    cpu>,  
                                num_devices=1)
```

Bases: `xfer.repurposer.Repurposer`

Base class for repurposers that extract features from layers in source neural network (Transfer) and train a meta-model using the extracted features (Learn).

Parameters

- **source_model** (`mxnet.mod.Module`) – Source neural network to do transfer learning from.
- **feature_layer_names** (`list[str]`) – Name of layer(s) in `source_model` from which features should be transferred.
- **context_function** (`function(int)->class:mx.context.Context`) – MXNet context function that provides device type context.
- **num_devices** (`int`) – Number of devices to use to extract features from `source_model`.

Methods

<code>__init__</code>	Initialize self.
<code>deserialize</code>	Abstract method.
<code>get_features_from_source_model</code>	Extract feature outputs from <code>feature_layer_names</code> in <code>source_model</code> , merge and return all features and labels.
<code>get_params</code>	Get parameters of repurposer that are in the constructor's argument list.
<code>predict_label</code>	Predict class labels on test data using the <code>target_model</code> (repurposed meta-model).
<code>predict_probability</code>	Predict class probabilities on test data using the <code>target_model</code> (repurposed meta-model).
<code>repurpose</code>	Train a meta-model using features extracted from training data through the source neural network.
<code>save_repurposer</code>	Serialize the repurposed model (<code>source_model</code> , <code>target_model</code> and supporting info) and save it to given <code>file_path</code> .
<code>serialize</code>	Abstract method.

Attributes

<code>feature_layer_names</code>	Names of the layers to extract features from.
<code>source_model</code>	Model to extract features from.

`get_params()`

Get parameters of repurposer that are in the constructor's argument list.

Return type `dict`

`feature_layer_names`

Names of the layers to extract features from.

`source_model`

Model to extract features from.

`repurpose(train_iterator: mxnet.io.io.DataIter)`

Train a meta-model using features extracted from training data through the source neural network.

Set `self.target_model` to the trained meta-model.

Parameters `train_iterator` – Training data iterator to use to extract features from `source_model`.

predict_probability (*test_iterator: mxnet.io.io.DataIter*)

Predict class probabilities on test data using the *target_model* (repurposed meta-model).

Parameters *test_iterator* (*mxnet.io.DataIter*) – Test data iterator to return predictions for.

Returns Predicted probabilities.

Return type `numpy.ndarray`

predict_label (*test_iterator: mxnet.io.io.DataIter*)

Predict class labels on test data using the *target_model* (repurposed meta-model).

Parameters *test_iterator* (*mxnet.io.DataIter*) – Test data iterator to return predictions for.

Returns Predicted labels.

Return type `numpy.ndarray`

get_features_from_source_model (*data_iterator: mxnet.io.io.DataIter*)

Extract feature outputs from *feature_layer_names* in *source_model*, merge and return all features and labels.

In addition, return mapping of *feature_layer_name* to indices in feature array.

Parameters *data_iterator* (*mxnet.io.DataIter*) – Iterator for data to be passed through the source network and extract features.

Returns features, *feature_indices_per_layer* and labels.

Return type `MetaModelData`

deserialize (*input_dict*)

Abstract method.

save_repurposer (*model_name, model_directory="", save_source_model=None*)

Serialize the repurposed model (*source_model*, *target_model* and supporting info) and save it to given *file_path*.

Parameters

- **model_name** (*str*) – Name to save repurposer to.
- **model_directory** (*str*) – File directory to save repurposer in.
- **save_source_model** (*boolean*) – Flag to choose whether to save repurposer source model. Will use default if set to None. (MetaModelRepurposer default: True, NeuralNetworkRepurposer default: False)

serialize (*file_prefix*)

Abstract method.

7.3 xfer.NeuralNetworkRepurposer

```
class xfer.NeuralNetworkRepurposer (source_model: mxnet.module.module.Module, context_function=<function cpu>, num_devices=1, batch_size=64, num_epochs=5, optimizer='sgd', optimizer_params=None)
```

Bases: `xfer.repurposer.Repurposer`

Base class for repurposers that create a target neural network from a source neural network through Transfer Learning.

- Transfer layer architecture and weights from a source neural network (Transfer) and
- Train a target neural network adapting the transferred network to new data set (Learn)

Parameters

- **source_model** (`mxnet.mod.Module`) – Source neural network to do transfer leaning from
- **context_function** (`function`) – MXNet context function that provides device type context
- **num_devices** (`int`) – Number of devices to use to train target neural network
- **batch_size** (`int`) – Size of data batches to be used for training the target neural network
- **num_epochs** (`int`) – Number of epochs to be used for training the target neural network
- **optimizer** (`str`) – Optimizer required by MXNet to train target neural network. Default: 'sgd'
- **optimizer_params** (`dict(str, float)`) – Optimizer params required by MXNet to train target neural network. Default: {'learning_rate': 1e-3}

Methods

<code>__init__</code>	Initialize self.
<code>deserialize</code>	Uses dictionary to set attributes of repurposer.
<code>get_params</code>	Get parameters of repurposer that are in the constructor
<code>predict_label</code>	Perform predictions on test data using the target_model (repurposed neural network).
<code>predict_probability</code>	Perform predictions on test data using the target_model (repurposed neural network).
<code>repurpose</code>	Train a neural network by transferring layers/weights from source_model.
<code>save_repurposer</code>	Serialize the repurposed model (source_model, target_model and supporting info) and save it to given file_path.
<code>serialize</code>	Serialize repurposer to dictionary.

`get_params()`

Get parameters of repurposer that are in the constructor

Return type `dict`

`repurpose(train_iterator: mxnet.io.io.DataIter)`

Train a neural network by transferring layers/weights from source_model. Set self.target_model to the repurposed neural network.

Parameters `train_iterator` (`mxnet.io.DataIter`) – Training data iterator to use to extract features from source_model

`predict_probability(test_iterator: mxnet.io.io.DataIter)`

Perform predictions on test data using the target_model (repurposed neural network).

Parameters `test_iterator` (`mxnet.io.DataIter`) – Test data iterator to return predictions for

Returns Predicted probabilities

Return type `numpy.ndarray`

predict_label (`test_iterator: mxnet.io.DataIter`)

Perform predictions on test data using the `target_model` (repurposed neural network).

Parameters `test_iterator` (`mxnet.io.DataIter`) – Test data iterator to return predictions for

Returns Predicted labels

Return type `numpy.ndarray`

serialize (`file_prefix`)

Serialize repurposer to dictionary.

Returns Dictionary describing repurposer

Return type `dict`

deserialize (`input_dict`)

Uses dictionary to set attributes of repurposer.

Parameters `input_dict` (`dict`) – Dictionary containing values for attributes to be set to

save_repurposer (`model_name, model_directory="", save_source_model=None`)

Serialize the repurposed model (`source_model`, `target_model` and supporting info) and save it to given `file_path`.

Parameters

- **model_name** (`str`) – Name to save repurposer to.
- **model_directory** (`str`) – File directory to save repurposer in.
- **save_source_model** (`boolean`) – Flag to choose whether to save repurposer source model. Will use default if set to `None`. (`MetaModelRepurposer` default: `True`, `NeuralNetworkRepurposer` default: `False`)

Repurposers:

<code>xfer.LrRepurposer</code>	Perform Transfer Learning through a Logistic Regression meta-model which repurposes the source neural network.
<code>xfer.SvmRepurposer</code>	Perform Transfer Learning through a Support Vector Machine (SVM) meta-model which repurposes the source neural network.
<code>xfer.GpRepurposer</code>	Repurpose source neural network to create a Gaussian Process (GP) meta-model through Transfer Learning.
<code>xfer.BnnRepurposer</code>	Perform Transfer Learning through a Bayesian Neural Network (BNN) meta-model which repurposes the source neural network.
<code>xfer.NeuralNetworkFineTuneRepurposer</code>	Class that creates a target neural network from a source neural network through Transfer Learning.
<code>xfer.NeuralNetworkRandomFreezeRepurposer</code>	Class that creates a target neural network from a source neural network through Transfer Learning.

7.4 xfer.LrRepurposer

class xfer.LrRepurposer(*source_model*: mxnet.module.module.Module, *feature_layer_names*, *context_function*=<function cpu>, *num_devices*=1, *tol*=0.0001, *c*=1.0, *n_jobs*=-1)

Bases: xfer.meta_model_repurposer.MetaModelRepurposer

Perform Transfer Learning through a Logistic Regression meta-model which repurposes the source neural network.

Parameters

- **source_model** (mxnet.mod.Module) – Source neural network to do Transfer Learning from.
- **feature_layer_names** (*list[str]*) – Name of layer(s) in source_model from which features should be transferred.
- **context_function** (function(int)->:class:mx.context.Context) – MXnet context function that provides device type context.
- **num_devices** (*int*) – Number of devices to use to extract features from source_model.
- **tol** (*float*) – Tolerance for stopping criteria.
- **c** (*float*) – Inverse of regularization strength; must be a positive float.
- **n_jobs** (*int*) – Number of CPU cores used during the cross-validation loop. If given a value of -1, all cores are used.

Methods

<code>__init__</code>	Initialize self.
<code>deserialize</code>	Uses dictionary to set attributes of repurposer.
<code>get_features_from_source_model</code>	Extract feature outputs from feature_layer_names in source_model, merge and return all features and labels.
<code>get_params</code>	Get parameters of repurposer that are in the constructor.
<code>predict_label</code>	Predict class labels on test data using the target_model (repurposed meta-model).
<code>predict_probability</code>	Predict class probabilities on test data using the target_model (repurposed meta-model).
<code>repurpose</code>	Train a meta-model using features extracted from training data through the source neural network.
<code>save_repurposer</code>	Serialize the repurposed model (source_model, target_model and supporting info) and save it to given file_path.
<code>serialize</code>	Saves repurposer (excluding source model) to file_prefix.json.

Attributes

<i>feature_layer_names</i>	Names of the layers to extract features from.
<i>source_model</i>	Model to extract features from.

get_params()

Get parameters of repurposer that are in the constructor.

Return type dict

serialize (*file_prefix*)

Saves repurposer (excluding source model) to file_prefix.json.

Parameters *file_prefix* (*str*) – Prefix to save file with.

deserialize (*input_dict*)

Uses dictionary to set attributes of repurposer.

Parameters *input_dict* (*dict*) – Dictionary containing values for attributes to be set to.

feature_layer_names

Names of the layers to extract features from.

get_features_from_source_model (*data_iterator: mxnet.io.io.DataIter*)

Extract feature outputs from feature_layer_names in source_model, merge and return all features and labels.

In addition, return mapping of feature_layer_name to indices in feature array.

Parameters *data_iterator* (*mxnet.io.DataIter*) – Iterator for data to be passed through the source network and extract features.

Returns features, feature_indices_per_layer and labels.

Return type MetaModelData

predict_label (*test_iterator: mxnet.io.io.DataIter*)

Predict class labels on test data using the target_model (repurposed meta-model).

Parameters *test_iterator* (*mxnet.io.DataIter*) – Test data iterator to return predictions for.

Returns Predicted labels.

Return type numpy.ndarray

predict_probability (*test_iterator: mxnet.io.io.DataIter*)

Predict class probabilities on test data using the target_model (repurposed meta-model).

Parameters *test_iterator* (*mxnet.io.DataIter*) – Test data iterator to return predictions for.

Returns Predicted probabilities.

Return type numpy.ndarray

repurpose (*train_iterator: mxnet.io.io.DataIter*)

Train a meta-model using features extracted from training data through the source neural network.

Set self.target_model to the trained meta-model.

Parameters *train_iterator* – Training data iterator to use to extract features from source_model.

save_repurposer (*model_name*, *model_directory*=", *save_source_model*=None)

Serialize the repurposed model (*source_model*, *target_model* and supporting info) and save it to given *file_path*.

Parameters

- **model_name** (*str*) – Name to save repurposer to.
- **model_directory** (*str*) – File directory to save repurposer in.
- **save_source_model** (*boolean*) – Flag to choose whether to save repurposer source model. Will use default if set to None. (MetaModelRepurposer default: True, NeuralNetworkRepurposer default: False)

source_model

Model to extract features from.

7.5 xfer.SvmRepurposer

class xfer.SvmRepurposer (*source_model*: mxnet.module.module.Module, *feature_layer_names*, *context_function*=<function cpu>, *num_devices*=1, *c*=1.0, *kernel*='linear', *gamma*='auto', *enable_probability_estimates*=False)

Bases: xfer.meta_model_repurposer.MetaModelRepurposer

Perform Transfer Learning through a Support Vector Machine (SVM) meta-model which repurposes the source neural network.

Parameters

- **source_model** (mxnet.mod.Module) – Source neural network to do transfer learning from.
- **feature_layer_names** (*list[str]*) – Name of layer(s) in *source_model* from which features should be transferred.
- **context_function** (function(int)->:class:mx.context.Context) – MXNet context function that provides device type context.
- **num_devices** (*int*) – Number of devices to use to extract features from *source_model*.
- **c** (*float*) – Penalty parameter C of the error term.
- **kernel** (*string*) – Specifies the kernel type to be used in the SVM algorithm in sklearn library. It must be one of 'linear', 'poly', 'rbf', 'sigmoid', 'precomputed' or a callable.
- **gamma** (*float*) – Kernel coefficient for 'rbf', 'poly' and 'sigmoid'. If gamma is 'auto' then 1/n_features will be used instead.
- **enable_probability_estimates** (*bool*) – Whether to enable probability estimates. This must be enabled for predict_probability to work and will slow down training.

Methods

<code>__init__</code>	Initialize self.
<code>deserialize</code>	Uses dictionary to set attributes of repurposer.

Continued on next page

Table 9 – continued from previous page

<i>get_features_from_source_model</i>	Extract feature outputs from <i>feature_layer_names</i> in <i>source_model</i> , merge and return all features and labels.
<i>get_params</i>	Get parameters of repurposer that are in the constructor.
<i>predict_label</i>	Predict class labels on test data using the <i>target_model</i> (repurposed meta-model).
<i>predict_probability</i>	Predict class probabilities on test data using the <i>target_model</i> (repurposed meta-model).
<i>repurpose</i>	Train a meta-model using features extracted from training data through the source neural network.
<i>save_repurposer</i>	Serialize the repurposed model (<i>source_model</i> , <i>target_model</i> and supporting info) and save it to given <i>file_path</i> .
<i>serialize</i>	Saves repurposer (excluding source model) to <i>file_prefix.json</i> .

Attributes

<i>feature_layer_names</i>	Names of the layers to extract features from.
<i>source_model</i>	Model to extract features from.

get_params()

Get parameters of repurposer that are in the constructor.

Return type `dict`

serialize(file_prefix)

Saves repurposer (excluding source model) to *file_prefix.json*.

Parameters *file_prefix* (`str`) – Prefix to save file with.

deserialize(input_dict)

Uses dictionary to set attributes of repurposer.

Parameters *input_dict* (`dict`) – Dictionary containing values for attributes to be set to.

feature_layer_names

Names of the layers to extract features from.

get_features_from_source_model (*data_iterator: mxnet.io.io.DataIter*)

Extract feature outputs from *feature_layer_names* in *source_model*, merge and return all features and labels.

In addition, return mapping of *feature_layer_name* to indices in feature array.

Parameters *data_iterator* (`mxnet.io.DataIter`) – Iterator for data to be passed through the source network and extract features.

Returns features, *feature_indices_per_layer* and labels.

Return type `MetaModelData`

predict_label (*test_iterator: mxnet.io.io.DataIter*)

Predict class labels on test data using the *target_model* (repurposed meta-model).

Parameters *test_iterator* (`mxnet.io.DataIter`) – Test data iterator to return predictions for.

Returns Predicted labels.

Return type `numpy.ndarray`

predict_probability (*test_iterator: mxnet.io.io.DataIter*)

Predict class probabilities on test data using the target_model (repurposed meta-model).

Parameters **test_iterator** (*mxnet.io.DataIter*) – Test data iterator to return predictions for.

Returns Predicted probabilities.

Return type `numpy.ndarray`

repurpose (*train_iterator: mxnet.io.io.DataIter*)

Train a meta-model using features extracted from training data through the source neural network.

Set self.target_model to the trained meta-model.

Parameters **train_iterator** – Training data iterator to use to extract features from source_model.

save_repurposer (*model_name, model_directory="", save_source_model=None*)

Serialize the repurposed model (source_model, target_model and supporting info) and save it to given file_path.

Parameters

- **model_name** (*str*) – Name to save repurposer to.
- **model_directory** (*str*) – File directory to save repurposer in.
- **save_source_model** (*boolean*) – Flag to choose whether to save repurposer source model. Will use default if set to None. (MetaModelRepurposer default: True, NeuralNetworkRepurposer default: False)

source_model

Model to extract features from.

7.6 xfer.GpRepurposer

```
class xfer.GpRepurposer(source_model: mxnet.module.module.Module, feature_layer_names, context_function=<function cpu>, num_devices=1, max_function_evaluations=100, apply_l2_norm=False)
```

Bases: `xfer.meta_model_repurposer.MetaModelRepurposer`

Repurpose source neural network to create a Gaussian Process (GP) meta-model through Transfer Learning.

Parameters

- **source_model** (*mxnet.mod.Module*) – Source neural network to do transfer learning from.
- **feature_layer_names** (*list[str]*) – Name of layer(s) in source_model from which features should be transferred.
- **context_function** (*function(int)->:class:mx.context.Context*) – MXNet context function that provides device type context.
- **num_devices** (*int*) – Number of devices to use to extract features from source_model.
- **max_function_evaluations** (*int*) – Maximum number of function evaluations to perform in GP optimization.

- **apply_l2_norm** (*bool*) – Whether to apply L2 normalization after extracting features from source neural network. If set to True, L2 normalization will be applied to features before passing to GP during training and prediction.

Methods

<code>__init__</code>	Initialize self.
<code>deserialize</code>	Uses dictionary to set attributes of repurposer.
<code>get_features_from_source_model</code>	Extract feature outputs from <code>feature_layer_names</code> in <code>source_model</code> , merge and return all features and labels.
<code>get_params</code>	Get parameters of repurposer that are in the constructor's argument list.
<code>predict_label</code>	Predict class labels on test data using the <code>target_model</code> (repurposed meta-model).
<code>predict_probability</code>	Predict class probabilities on test data using the <code>target_model</code> (repurposed meta-model).
<code>repurpose</code>	Train a meta-model using features extracted from training data through the source neural network.
<code>save_repurposer</code>	Serialize the repurposed model (<code>source_model</code> , <code>target_model</code> and supporting info) and save it to given <code>file_path</code> .
<code>serialize</code>	Serialize the GP repurposer (model and supporting info) and save to file.

Attributes

<code>feature_layer_names</code>	Names of the layers to extract features from.
<code>source_model</code>	Model to extract features from.

`get_params()`

Get parameters of repurposer that are in the constructor's argument list.

Return type `dict`

`serialize(file_prefix)`

Serialize the GP repurposer (model and supporting info) and save to file.

Parameters `file_prefix` (*str*) – Prefix of file path to save the serialized repurposer to.

Returns `None`

`deserialize(input_dict)`

Uses dictionary to set attributes of repurposer.

Parameters `input_dict` (*dict*) – Dictionary containing values for attributes to be set to.

Returns `None`

`feature_layer_names`

Names of the layers to extract features from.

`get_features_from_source_model(data_iterator: mxnet.io.io.DataIter)`

Extract feature outputs from `feature_layer_names` in `source_model`, merge and return all features and labels.

In addition, return mapping of `feature_layer_name` to indices in feature array.

Parameters `data_iterator` (`mxnet.io.DataIter`) – Iterator for data to be passed through the source network and extract features.

Returns features, `feature_indices_per_layer` and labels.

Return type `MetaModelData`

predict_label (*test_iterator: mxnet.io.io.DataIter*)

Predict class labels on test data using the `target_model` (repurposed meta-model).

Parameters `test_iterator` (`mxnet.io.DataIter`) – Test data iterator to return predictions for.

Returns Predicted labels.

Return type `numpy.ndarray`

predict_probability (*test_iterator: mxnet.io.io.DataIter*)

Predict class probabilities on test data using the `target_model` (repurposed meta-model).

Parameters `test_iterator` (`mxnet.io.DataIter`) – Test data iterator to return predictions for.

Returns Predicted probabilities.

Return type `numpy.ndarray`

repurpose (*train_iterator: mxnet.io.io.DataIter*)

Train a meta-model using features extracted from training data through the source neural network.

Set `self.target_model` to the trained meta-model.

Parameters `train_iterator` – Training data iterator to use to extract features from `source_model`.

save_repurposer (*model_name, model_directory="", save_source_model=None*)

Serialize the repurposed model (`source_model`, `target_model` and supporting info) and save it to given `file_path`.

Parameters

- **model_name** (*str*) – Name to save repurposer to.
- **model_directory** (*str*) – File directory to save repurposer in.
- **save_source_model** (*boolean*) – Flag to choose whether to save repurposer source model. Will use default if set to `None`. (`MetaModelRepurposer` default: `True`, `NeuralNetworkRepurposer` default: `False`)

source_model

Model to extract features from.

7.7 xfer.BnnRepurposer

```
class xfer.BnnRepurposer (source_model:      mxnet.module.module.Module, feature_layer_names,
                          context_function=<function      cpu>,          num_devices=1,
                          bnn_context_function=<function  cpu>, sigma=100.0, num_layers=1,
                          n_hidden=10,          num_samples_mc=3,          learning_rate=0.001,
                          batch_size=20,          num_epochs=200,          start_annealing=None,
                          end_annealing=None, num_samples_mc_prediction=100, verbose=0)
```

Bases: xfer.meta_model_repurposer.MetaModelRepurposer

Perform Transfer Learning through a Bayesian Neural Network (BNN) meta-model which repurposes the source neural network.

Parameters

- **source_model** (mxnet.mod.Module) – Source neural network to do transfer learning from.
- **feature_layer_names** (*list[str]*) – Name of layer(s) in source_model from which features should be transferred.
- **context_function** (function(int)->:class:mx.context.Context) – MXNet context function that provides device type context. It is used to extract features with the source model.
- **num_devices** (*int*) – Number of devices to use to extract features from source_model.
- **bnn_context_function** (function(int)->:class:mx.context.Context) – MXNet context function used to train the BNN.
- **sigma** (*float*) – Standard deviation of the Gaussian prior used for the weights of the BNN meta-model ($w_i \sim N(0, \sigma^2)$).
- **num_layers** (*int*) – Number of layers of the BNN meta-model.
- **n_hidden** (*int*) – Dimensionality of the hidden layers of the BNN meta-model (all hidden layers have the same dimensionality).
- **num_samples_mc** (*int*) – Number of samples used for the Monte Carlo approximation of the variational bound.
- **learning_rate** (*float*) – Learning rate for the BNN meta-model training.
- **batch_size** (*int*) – Mini-batch size for the BNN meta-model training.
- **num_epochs** (*int*) – Number of epochs for the BNN meta-model training.
- **start_annealing** (*int*) – To help the training of the BNN meta-model, we anneal the KL term using a weight that varies from zero to one. start_annealing determines the epoch at which the annealing weight start to increase linearly until it reaches one in the epoch given by end_annealing.
- **end_annealing** (*int*) – Determines the epoch at which the annealing process of the KL term ends.
- **step_annealing_sample_weight** (*float*) – Amount that the annealing weight is incremented in each epoch (from start_annealing to end_annealing).
- **num_samples_mc_prediction** (*int*) – Number of Monte Carlo samples to use on prediction.
- **verbose** (*bool*) – Flag to control whether accuracy monitoring is logged during repurposing.

- **annealing_weight** (*float*) – Annealing weight in the current epoch.
- **train_acc** (*list[float]*) – Accuracy in training set in each epoch.
- **test_acc** (*list[float]*) – Accuracy in validation set in each epoch.
- **moving_loss_total** (*list[float]*) – Total loss (negative ELBO) smoothed across epochs.
- **average_loss** (*list[float]*) – Average loss (negative ELBO) per data point.
- **anneal_weights** (*list[float]*) – Annealing weight used in each epoch.

Methods

<code>__init__</code>	Initialize self.
<code>deserialize</code>	Uses dictionary to set attributes of repurposer.
<code>get_features_from_source_model</code>	Extract feature outputs from feature_layer_names in source_model, merge and return all features and labels.
<code>get_params</code>	Get parameters of repurposer that are in the constructor.
<code>predict_label</code>	Predict class labels on test data using the target_model (repurposed meta-model).
<code>predict_probability</code>	Predict class probabilities on test data using the target_model (repurposed meta-model).
<code>repurpose</code>	Train a meta-model using features extracted from training data through the source neural network.
<code>save_repurposer</code>	Serialize the repurposed model (source_model, target_model and supporting info) and save it to given file_path.
<code>serialize</code>	Saves repurposer (excluding source model) to file_prefix.json, file_prefix_posterior.json, file_prefix_posterior_params.npz.

Attributes

<code>end_annealing</code>	Determines the epoch at which the annealing process of the KL term ends.
<code>feature_layer_names</code>	Names of the layers to extract features from.
<code>source_model</code>	Model to extract features from.
<code>start_annealing</code>	Determines the epoch at which the annealing process of the KL term starts.

get_params()

Get parameters of repurposer that are in the constructor.

Return type dict

start_annealing

Determines the epoch at which the annealing process of the KL term starts.

end_annealing

Determines the epoch at which the annealing process of the KL term ends.

feature_layer_names

Names of the layers to extract features from.

get_features_from_source_model (*data_iterator: mxnet.io.io.DataIter*)

Extract feature outputs from `feature_layer_names` in `source_model`, merge and return all features and labels.

In addition, return mapping of `feature_layer_name` to indices in feature array.

Parameters `data_iterator` (*mxnet.io.DataIter*) – Iterator for data to be passed through the source network and extract features.

Returns features, `feature_indices_per_layer` and labels.

Return type `MetaModelData`

predict_label (*test_iterator: mxnet.io.io.DataIter*)

Predict class labels on test data using the `target_model` (repurposed meta-model).

Parameters `test_iterator` (*mxnet.io.DataIter*) – Test data iterator to return predictions for.

Returns Predicted labels.

Return type `numpy.ndarray`

predict_probability (*test_iterator: mxnet.io.io.DataIter*)

Predict class probabilities on test data using the `target_model` (repurposed meta-model).

Parameters `test_iterator` (*mxnet.io.DataIter*) – Test data iterator to return predictions for.

Returns Predicted probabilities.

Return type `numpy.ndarray`

repurpose (*train_iterator: mxnet.io.io.DataIter*)

Train a meta-model using features extracted from training data through the source neural network.

Set `self.target_model` to the trained meta-model.

Parameters `train_iterator` – Training data iterator to use to extract features from `source_model`.

save_repurposer (*model_name, model_directory="", save_source_model=None*)

Serialize the repurposed model (`source_model`, `target_model` and supporting info) and save it to given `file_path`.

Parameters

- **model_name** (*str*) – Name to save repurposer to.
- **model_directory** (*str*) – File directory to save repurposer in.
- **save_source_model** (*boolean*) – Flag to choose whether to save repurposer source model. Will use default if set to `None`. (`MetaModelRepurposer` default: `True`, `NeuralNetworkRepurposer` default: `False`)

source_model

Model to extract features from.

serialize (*file_prefix*)

Saves repurposer (excluding `source_model`) to `file_prefix.json`, `file_prefix_posterior.json`, `file_prefix_posterior_params.npz`.

Parameters `file_prefix` (*str*) – Prefix to save file with.

deserialize (*input_dict*)

Uses dictionary to set attributes of repurposer.

Parameters *input_dict* (*dict*) – Dictionary containing values for attributes to be set to.

7.8 xfer.NeuralNetworkFineTuneRepurposer

```
class xfer.NeuralNetworkFineTuneRepurposer (source_model: mxnet.module.module.Module,
                                             transfer_layer_name,      target_class_count,
                                             context_function=<function      cpu>,
                                             num_devices=1,           batch_size=64,
                                             num_epochs=5,      optimizer='sgd',    opti-
                                             mizer_params=None)
```

Bases: xfer.neural_network_repurposer.NeuralNetworkRepurposer

Class that creates a target neural network from a source neural network through Transfer Learning. It transfers layers from source model and fine-tunes the network to learn from target data set.

Steps involved:

- Layers and weights of source model are transferred from input layer up to and including layer 'transfer_layer_name'
- A fully connected layer with nodes equal to number of classes in target data set is added after the transfer layer
- A softmax layer is added on top of the new fully connected layer
- This modified network represents the target model and is fine tuned to adapt weights to the target data set

Parameters

- **source_model** (*mxnet.mod.Module*) – Source neural network to do transfer leaning from. Can be None in a predict-only case
- **transfer_layer_name** (*str*) – Name of layer up to which layers/weights in source_model need to be transferred. For example, name of layer before the last fully connected layer in source network. Can be None in a predict-only case
- **target_class_count** (*int*) – Number of classes to train the target neural network for. Can be None in a predict-only case
- **context_function** (*function*) – MXNet context function that provides device type context
- **num_devices** (*int*) – Number of devices to use to train target neural network
- **batch_size** (*int*) – Size of data batches to be used for training the target neural network
- **num_epochs** (*int*) – Number of epochs to be used for training the target neural network
- **optimizer** (*str*) – Optimizer required by MXNet to train target neural network. Default: 'sgd'
- **optimizer_params** (*dict(str, float)*) – Optimizer params required by MXNet to train target neural network. Default: {'learning_rate': 1e-3}

Methods

<code>__init__</code>	Initialize self.
<code>deserialize</code>	Uses dictionary to set attributes of repurposer.
<code>get_params</code>	Get parameters of repurposer that are in the constructor
<code>predict_label</code>	Perform predictions on test data using the target_model (repurposed neural network).
<code>predict_probability</code>	Perform predictions on test data using the target_model (repurposed neural network).
<code>repurpose</code>	Train a neural network by transferring layers/weights from source_model.
<code>save_repurposer</code>	Serialize the repurposed model (source_model, target_model and supporting info) and save it to given file_path.
<code>serialize</code>	Serialize repurposer to dictionary.

get_params()

Get parameters of repurposer that are in the constructor

Return type dict

deserialize(input_dict)

Uses dictionary to set attributes of repurposer.

Parameters input_dict (dict) – Dictionary containing values for attributes to be set to

predict_label(test_iterator: mxnet.io.io.DataIter)

Perform predictions on test data using the target_model (repurposed neural network).

Parameters test_iterator (mxnet.io.DataIter) – Test data iterator to return predictions for

Returns Predicted labels

Return type numpy.ndarray

predict_probability(test_iterator: mxnet.io.io.DataIter)

Perform predictions on test data using the target_model (repurposed neural network).

Parameters test_iterator (mxnet.io.DataIter) – Test data iterator to return predictions for

Returns Predicted probabilities

Return type numpy.ndarray

repurpose(train_iterator: mxnet.io.io.DataIter)

Train a neural network by transferring layers/weights from source_model. Set self.target_model to the repurposed neural network.

Parameters train_iterator (mxnet.io.DataIter) – Training data iterator to use to extract features from source_model

save_repurposer(model_name, model_directory="", save_source_model=None)

Serialize the repurposed model (source_model, target_model and supporting info) and save it to given file_path.

Parameters

- **model_name (str)** – Name to save repurposer to.
- **model_directory (str)** – File directory to save repurposer in.

- **save_source_model** (*boolean*) – Flag to choose whether to save repurposer source model. Will use default if set to None. (MetaModelRepurposer default: True, NeuralNetworkRepurposer default: False)

serialize (*file_prefix*)

Serialize repurposer to dictionary.

Returns Dictionary describing repurposer

Return type `dict`

7.9 xfer.NeuralNetworkRandomFreezeRepurposer

```
class xfer.NeuralNetworkRandomFreezeRepurposer (source_model:
                                                mxnet.module.module.Module,      tar-
                                                get_class_count,  fixed_layers,  ran-
                                                dom_layers,    num_layers_to_drop=2,
                                                context_function=<function      cpu>,
                                                num_devices=1,      batch_size=64,
                                                num_epochs=5,      optimizer='sgd',
                                                optimizer_params=None)
```

Bases: `xfer.neural_network_repurposer.NeuralNetworkRepurposer`

Class that creates a target neural network from a source neural network through Transfer Learning. It transfers layers from source model and fine-tunes the network to learn from target data set.

Steps involved:

- Layers and weights of source model are transferred
- N layers are removed from the output of the model
- A fully connected layer with nodes equal to number of classes in target data set is added to the model output
- A softmax layer is added on top of the new fully connected layer
- Some layers are frozen and some randomly initialized and then model is fine-tuned

Parameters

- **source_model** (`mxnet.mod.Module`) – Source neural network to do transfer leaning from. Can be None in a predict-only case
- **target_class_count** (*int*) – Number of classes to train the target neural network for. Can be None in a predict-only case
- **fixed_layers** (*list[str]*) – List of layers to keep weights frozen for
- **random_layers** (*list[str]*) – List of layers to randomly reinitialise
- **num_layers_to_drop** (*int*) – Number of layers to remove from model output
- **context_function** (*function*) – MXNet context function that provides device type context
- **num_devices** (*int*) – Number of devices to use to train target neural network
- **batch_size** (*int*) – Size of data batches to be used for training the target neural network
- **num_epochs** (*int*) – Number of epochs to be used for training the target neural network

- **optimizer** (*str*) – Optimizer required by MXNet to train target neural network. Default: 'sgd'
- **optimizer_params** (*dict(str, float)*) – Optimizer params required by MXNet to train target neural network. Default: {'learning_rate': 1e-3}

Methods

<code>__init__</code>	Initialize self.
<code>deserialize</code>	Uses dictionary to set attributes of repurposer.
<code>get_params</code>	Get parameters of repurposer that are in the constructor
<code>predict_label</code>	Perform predictions on test data using the target_model (repurposed neural network).
<code>predict_probability</code>	Perform predictions on test data using the target_model (repurposed neural network).
<code>repurpose</code>	Train a neural network by transferring layers/weights from source_model.
<code>save_repurposer</code>	Serialize the repurposed model (source_model, target_model and supporting info) and save it to given file_path.
<code>serialize</code>	Serialize repurposer to dictionary.

get_params()

Get parameters of repurposer that are in the constructor

Return type `dict`

deserialize(input_dict)

Uses dictionary to set attributes of repurposer.

Parameters **input_dict** (*dict*) – Dictionary containing values for attributes to be set to

predict_label(test_iterator: mxnet.io.io.DataIter)

Perform predictions on test data using the target_model (repurposed neural network).

Parameters **test_iterator** (*mxnet.io.DataIter*) – Test data iterator to return predictions for

Returns Predicted labels

Return type `numpy.ndarray`

predict_probability(test_iterator: mxnet.io.io.DataIter)

Perform predictions on test data using the target_model (repurposed neural network).

Parameters **test_iterator** (*mxnet.io.DataIter*) – Test data iterator to return predictions for

Returns Predicted probabilities

Return type `numpy.ndarray`

repurpose(train_iterator: mxnet.io.io.DataIter)

Train a neural network by transferring layers/weights from source_model. Set self.target_model to the repurposed neural network.

Parameters **train_iterator** (*mxnet.io.DataIter*) – Training data iterator to use to extract features from source_model

save_repurposer (*model_name*, *model_directory*=", *save_source_model*=None)

Serialize the repurposed model (source_model, target_model and supporting info) and save it to given file_path.

Parameters

- **model_name** (*str*) – Name to save repurposer to.
- **model_directory** (*str*) – File directory to save repurposer in.
- **save_source_model** (*boolean*) – Flag to choose whether to save repurposer source model. Will use default if set to None. (MetaModelRepurposer default: True, NeuralNetworkRepurposer default: False)

serialize (*file_prefix*)

Serialize repurposer to dictionary.

Returns Dictionary describing repurposer

Return type `dict`

Model Handler

<code>xfer.model_handler.ModelHandler</code>	Class for model manipulation and feature extraction.
<code>xfer.model_handler.exceptions</code>	Exceptions for Model Handler.
<code>xfer.model_handler.consts</code>	Model Handler constants.

8.1 xfer.model_handler.ModelHandler

class xfer.model_handler.**ModelHandler** (*module*, *context_function=<function cpu>*, *num_devices=1*, *data_name='data'*)

Bases: `object`

Class for model manipulation and feature extraction.

Parameters

- **module** (`mx.module.Module`) – MXNet module to be manipulated.
- **context_function** (*function*) – MXNet context function.
- **num_devices** (*int*) – Number of devices to run process on.
- **data_name** (*str*) – Name of input layer of model.

Methods

<code>__init__</code>	Initialize self.
<code>add_layer_bottom</code>	Add layer to input of model.
<code>add_layer_top</code>	Add layer to output of model.
<code>drop_layer_bottom</code>	Remove layers from input of model.
<code>drop_layer_top</code>	Remove layers from output of model.
<code>get_layer_names_matching_type</code>	Return names of layers of specified type.

Continued on next page

Table 2 – continued from previous page

<code>get_layer_output</code>	Function to extract features from data iterator with model.
<code>get_layer_parameters</code>	Get list of layer parameters associated with the the layer names given.
<code>get_layer_type</code>	Return type of named layer.
<code>get_module</code>	Return MXNet Module using the model symbol and parameters.
<code>save_symbol</code>	Serialise model symbol graph.
<code>update_sym</code>	Update symbol attribute, layer names, and layer types dict and clean parameters.
<code>visualize_net</code>	Display computational graph of model.

Attributes

<code>layer_names</code>	Get list of names of model layers.
--------------------------	------------------------------------

drop_layer_top (*num_layers_to_drop=1*)

Remove layers from output of model.

Parameters *n* (*int*) – Number of layers to remove from model output.

drop_layer_bottom (*num_layers_to_drop=1*)

Remove layers from input of model.

Parameters *n* (*int*) – Number of layers to remove from model input.

add_layer_top (*layer_list*)

Add layer to output of model. model layers = (layer1, layer2, layer3), layer_list = [layerA, layerB] -> model layers = (layer1, layer2, layer3, layerA, layerB)

Parameters **layer_list** (*list(mx.symbol)*) – List of MxNet symbol layers to be added to model output.

add_layer_bottom (*layer_list*)

Add layer to input of model. model layers = (layer1, layer2, layer3), layer_list = [layerA, layerB] -> model layers = (layerA, layerB, layer1, layer2, layer3)

Parameters **layer_list** (*list(mx.symbol)*) – List of MxNet symbol layers to be added to model input.

get_module (*iterator, fixed_layer_parameters=None, random_layer_parameters=None*)

Return MXNet Module using the model symbol and parameters.

Parameters

- **iterator** (*mxnet.io.DataIter*) – MXNet iterator to be used with model.
- **fixed_layer_parameters** (*list(str)*) – List of layer parameters to keep fixed.
- **random_layer_parameters** (*list(str)*) – List of layer parameters to randomise.

Returns MXNet module

Return type *mx.module.Module*

get_layer_type (*layer_name*)

Return type of named layer.

Parameters **name** (*str*) – Name of layer being inspected.

Returns Layer type

Return type *str*

get_layer_names_matching_type (*layer_type*)

Return names of layers of specified type.

Parameters **layer_type** (*str*) – Return list of layers of this type.

Returns Names of layers with specified type

Return type *list(str)*

get_layer_output (*data_iterator, layer_names*)

Function to extract features from data iterator with model. Returns a dictionary of layer_name -> numpy array of features extracted (flattened).

Parameters

- **data_iterator** (*mxnet.io.DataIter*) – Iterator containing input data.
- **layer_names** (*list(str)*) – List of names of layers to extract features from.

Returns Ordered Dictionary of features ({layer_name: features}), list of labels Layer names in the ordered dictionary follow the same order as input list of layer_names

Return type *OrderedDict[str, numpy.array], list(int)*

get_layer_parameters (*layer_names*)

Get list of layer parameters associated with the the layer names given.

Parameters **layer_names** (*list(str)*) – List of layer names.

Returns List of layer parameters

Return type *list(str)*

visualize_net ()

Display computational graph of model.

save_symbol (*model_name*)

Serialise model symbol graph.

Parameters **model_name** (*str*) – Prefix to file name (model_name-symbol.json).

layer_names

Get list of names of model layers.

Returns List of layer names

Return type *list[str]*

update_sym (*new_symbol*)

Update symbol attribute, layer names, and layer types dict and clean parameters.

Parameters **new_symbol** (*mx.symbol.Symbol*) – Symbol with which to update Model-Handler.

8.2 xfer.model_handler.exceptions

Exceptions for Model Handler.

Exceptions

<i>ModelArchitectureError</i>	Exception type for errors caused when model architecture is incorrect or mismatched
<i>ModelError</i>	Exception type for errors caused by invalid model actions.

exception xfer.model_handler.exceptions.**ModelError**

Bases: `Exception`

Exception type for errors caused by invalid model actions.

args

with_traceback()

Exception.with_traceback(tb) – set self.__traceback__ to tb and return self.

exception xfer.model_handler.exceptions.**ModelArchitectureError**

Bases: `Exception`

Exception type for errors caused when model architecture is incorrect or mismatched

args

with_traceback()

Exception.with_traceback(tb) – set self.__traceback__ to tb and return self.

8.3 xfer.model_handler.consts

Model Handler constants.

Classes

<i>LayerType</i>	An enumeration.
------------------	-----------------

class xfer.model_handler.consts.**LayerType**

Bases: `enum.Enum`

An enumeration.

FULLYCONNECTED = 'FullyConnected'

CONVOLUTION = 'Convolution'

ACTIVATION = 'Activation'

POOLING = 'Pooling'

FLATTEN = 'Flatten'

SOFTMAXOUTPUT = 'SoftmaxOutput'

DROPOUT = 'Dropout'

CONCAT = 'Concat'

BATCHNORM = 'BatchNorm'

```
PLUS = '_Plus'  
SVMOUTPUT = 'SVMOutput'  
LRN = 'LRN'  
DATA = 'Data'  
MEAN = 'mean'  
EMBEDDING = 'Embedding'
```

Writing a custom Repurposer

Xfer implements and supports two kinds of Repurposers:

- **Meta-model Repurposer** - this uses the source model to extract features and then fits a meta-model to the features
- **Neural network Repurposer** - this modifies the source model to create a target model

Below are examples of creating custom Repurposers for both classes

9.1 Setup

First import relevant modules, define data iterators and load a source model

```
In [1]: import warnings
        warnings.filterwarnings("ignore")

        import logging
        logging.disable(logging.WARNING)

        import xfer

        import os
        import glob
        import mxnet as mx
        import random
        from sklearn.model_selection import StratifiedShuffleSplit
        from sklearn.metrics import classification_report

        random.seed(1)

In [2]: def get_iterators_from_folder(data_dir, train_size=0.6, batchsize=10, label_name='softmax_label')
        """
        Method to create iterators from data stored in a folder with the following structure:
        /data_dir
          /class1
```

```

        class1_img1
        class1_img2
        ...
        class1_imgN
    /class2
        class2_img1
        class2_img2
        ...
        class2_imgN
    ...
    /classN
"""
# assert dir exists
if not os.path.isdir(data_dir):
    raise ValueError('Directory not found: {}'.format(data_dir))
# get class names
classes = [x.split('/')[0] for x in glob.glob(data_dir+'/*')]
classes.sort()
fnames = []
labels = []
for c in classes:
    # get all the image filenames and labels
    images = glob.glob(data_dir+'/' + c + '/*')
    images.sort()
    fnames += images
    labels += [c]*len(images)
# create label2id mapping
id2label = dict(enumerate(set(labels)))
label2id = dict((v,k) for k, v in id2label.items())

# get indices of train and test
sss = StratifiedShuffleSplit(n_splits=2, test_size=None, train_size=train_size, random_state=None)
train_indices, test_indices = next(sss.split(labels, labels))

train_img_list = []
test_img_list = []
train_labels = []
test_labels = []
# create imglist for training and test
for idx in train_indices:
    train_img_list.append([label2id[labels[idx]], fnames[idx]])
    train_labels.append(label2id[labels[idx]])
for idx in test_indices:
    test_img_list.append([label2id[labels[idx]], fnames[idx]])
    test_labels.append(label2id[labels[idx]])

# make iterators
train_iterator = mx.image.ImageIter(batchsize, (3,224,224), imglist=train_img_list, label_paths=train_labels,
                                     path_root='')
test_iterator = mx.image.ImageIter(batchsize, (3,224,224), imglist=test_img_list, label_paths=test_labels,
                                   path_root='')

return train_iterator, test_iterator, train_labels, test_labels, id2label, label2id
In [3]: dataset = 'test_images' # options are: 'test_sketches', 'test_images_sketch', 'mnist-50', 'test_images'
        num_classes = 4

        train_iterator, test_iterator, train_labels, test_labels, id2label, label2id = get_iterators(dataset, num_classes)
In [4]: # Download vgg19 (trained on imagenet)

```

```
path = 'http://data.mxnet.io/models/imagenet/'
[mx.test_utils.download(path+'vgg/vgg19-0000.params'),
mx.test_utils.download(path+'vgg/vgg19-symbol.json')]
```

```
Out[4]: ['vgg19-0000.params', 'vgg19-symbol.json']
```

```
In [5]: # This will be the source model we use for repurposing later
source_model = mx.module.Module.load('vgg19', 0, label_names=['prob_label'])
```

9.2 Custom Meta-model Repurposer

We will create a new Repurposer that uses the KNN algorithm as a meta-model. The resulting Meta-model Repurposer will classify the features extracted by the neural network source model.

```
In [6]: from sklearn.neighbors import KNeighborsClassifier
```

9.2.1 Definition

```
In [7]: class KNNRepurposer(xfer.MetaModelRepurposer):
    def __init__(self, source_model: mx.mod.Module, feature_layer_names, context_function=mx.Context,
                 n_neighbors=5, weights='uniform', algorithm='auto', leaf_size=30, p=2, metric='l2'):
        # Call init() of parent
        super(KNNRepurposer, self).__init__(source_model, feature_layer_names, context_function)

        # Initialise parameters specific to the KNN algorithm
        self.n_neighbors = n_neighbors
        self.weights = weights
        self.algorithm = algorithm
        self.leaf_size = leaf_size
        self.p = p
        self.metric = metric
        self.metric_params = metric_params
        self.n_jobs = n_jobs

    # Define function that takes a set of features and labels and returns a trained model.
    # feature_indices_per_layer is a dictionary which gives the feature indices which correspond
    # to each layer's features.
    def _train_model_from_features(self, features, labels, feature_indices_per_layer=None):
        lin_model = KNeighborsClassifier(n_neighbors=self.n_neighbors,
                                       weights=self.weights,
                                       algorithm=self.algorithm,
                                       leaf_size=self.leaf_size,
                                       p=self.p,
                                       metric=self.metric,
                                       metric_params=self.metric_params)

        lin_model.fit(features, labels)
        return lin_model

    # Define a function that predicts the class probability given features
    def _predict_probability_from_features(self, features):
        return self.target_model.predict_proba(features)

    # Define a function that predicts the class label given features
    def _predict_label_from_features(self, features):
        return self.target_model.predict(features)

    # In order to make your repurposer serialisable, you will need to implement functions
```

```
# which convert your model's parameters to a dictionary.
def get_params(self):
    """
    This function should return a dictionary of all the parameters of the repurposer that
    are in the repurposer constructor arguments.
    """
    param_dict = super().get_params()
    param_dict['n_neighbors'] = self.n_neighbors
    param_dict['weights'] = self.weights
    param_dict['algorithm'] = self.algorithm
    param_dict['leaf_size'] = self.leaf_size
    param_dict['p'] = self.p
    param_dict['metric'] = self.metric
    param_dict['metric_params'] = self.metric_params
    param_dict['n_jobs'] = self.n_jobs
    return param_dict

# Some repurposers will need a get_attributes() and set_attributes() to get and set the p
# of the repurposer that are not in the constructor argument. An example is shown below:

# def get_attributes(self):
#     """
#     This function should return a dictionary of all the parameters of the repurposer th
#     are NOT in the constructor arguments.
#     This function does not need to be defined if the repurposer has no specific attrib
#     """
#     param_dict = super().get_attributes()
#     param_dict['example_attribute'] = self.example_attribute
#     return param_dict

# def set_attributes(self, input_dict):
#     super().set_attributes(input_dict)
#     self.example_attribute = input_dict['example_attribute']

def serialize(self, file_prefix):
    """
    Saves repurposer (excluding source model) to file_prefix.json.
    This method converts the repurposer to dictionary and saves as a json.

    :param str file_prefix: Prefix to save file with
    """
    output_dict = {}
    output_dict[repurposer_keys.PARAMS] = self.get_params()
    output_dict[repurposer_keys.TARGET_MODEL] = target_model_to_dict() # This should be
    output_dict.update(self.get_attributes())

    utils.save_json(file_prefix, output_dict)

def deserialize(self, input_dict):
    """
    Uses dictionary to set attributes of repurposer

    :param dict input_dict: Dictionary containing values for attributes to be set to
    """
    self.set_attributes(input_dict) # Set attributes of the repurposer from input_dict
    self.target_model = target_model_from_dict() # Unpack dictionary representation of t
```

9.2.2 Use

```
In [8]: repurposerKNN = KNNRepurposer(source_model, ['fc8'])
In [9]: repurposerKNN.repurpose(train_iterator)
In [10]: results = repurposerKNN.predict_label(test_iterator)
In [11]: print(classification_report(y_pred=results, y_true=test_labels))
```

	precision	recall	f1-score	support
0	1.00	0.50	0.67	2
1	0.67	1.00	0.80	2
2	1.00	1.00	1.00	2
3	1.00	1.00	1.00	2
avg / total	0.92	0.88	0.87	8

9.3 Custom Neural Network Repurposer

Now we will define a custom Neural Network Repurposer which performs transfer learning by:

1. taking the original source neural network and keeping all layers up to `transfer_layer_name`
2. adding two fully connected layers on the top
3. fine-tuning with any conv layers frozen

9.3.1 Definition

```
In [12]: class Add2FullyConnectedRepurposer(xfer.NeuralNetworkRepurposer):
    def __init__(self, source_model: mx.mod.Module, transfer_layer_name, num_nodes, target_class_count,
                 context_function=mx.context.cpu, num_devices=1, batch_size=64, num_epochs=100):
        super().__init__(source_model, context_function, num_devices, batch_size, num_epochs)

        # initialise parameters
        self.transfer_layer_name = transfer_layer_name
        self.num_nodes = num_nodes
        self.target_class_count = target_class_count

    def _get_target_symbol(self, source_model_layer_names):
        # Check if 'transfer_layer_name' is present in source model
        if self.transfer_layer_name not in source_model_layer_names:
            raise ValueError('transfer_layer_name: {} not found in source model'.format(self.transfer_layer_name))

        # Create target symbol by transferring layers from source model up to 'transfer_layer_name'
        transfer_layer_key = self.transfer_layer_name + '_output' # layer key with output symbol
        source_symbol = self.source_model.symbol.get_internals()
        target_symbol = source_symbol[transfer_layer_key]
        return target_symbol

    # All Neural Network Repurposers must implement this function which takes a training iterator
    def _create_target_module(self, train_iterator: mx.io.DataIter):
        # Create model handler to manipulate the source model
        model_handler = xfer.model_handler.ModelHandler(self.source_model, self.context_function)
```

```
# Create target symbol by transferring layers from source model up to and including
target_symbol = self._get_target_symbol(model_handler.layer_names)

# Update model handler by replacing source symbol with target symbol
# and cleaning up weights of layers that were not transferred
model_handler.update_sym(target_symbol)

# Add a fully connected layer (with nodes equal to number of target classes) and a s
fully_connected_layer1 = mx.sym.FullyConnected(num_hidden=self.num_nodes, name='fc_1
fully_connected_layer2 = mx.sym.FullyConnected(num_hidden=self.target_class_count, n
softmax_output_layer = mx.sym.SoftmaxOutput(name=train_iterator.provide_label[0][0]
model_handler.add_layer_top([fully_connected_layer1, fully_connected_layer2, softm

# Get fixed layers
conv_layer_names = model_handler.get_layer_names_matching_type('Convolution')
conv_layer_params = model_handler.get_layer_parameters(conv_layer_names)

# Create and return target mxnet module using the new symbol and params
return model_handler.get_module(train_iterator, fixed_layer_parameters=conv_layer_p

# To be serialisable, Neural Network Repurposers require get_params, get_attributes, set
```

9.3.2 Use

```
In [13]: # instantiate repurposer
         repurposer2Fc = Add2FullyConnectedRepurposer(source_model, transfer_layer_name='fc7', num_n

In [14]: train_iterator.reset()
         repurposer2Fc.repurpose(train_iterator)

In [15]: results = repurposer2Fc.predict_label(test_iterator)

In [16]: print(classification_report(y_pred=results, y_true=test_labels))
```

	precision	recall	f1-score	support
0	1.00	0.50	0.67	2
1	1.00	1.00	1.00	2
2	1.00	1.00	1.00	2
3	0.67	1.00	0.80	2
avg / total	0.92	0.88	0.87	8

CHAPTER 10

Indices and tables

- `genindex`
- `modindex`
- `search`

X

`xfer.model_handler.consts`, [72](#)

`xfer.model_handler.exceptions`, [71](#)

A

ACTIVATION (xfer.model_handler.consts.LayerType attribute), 72
 add_layer_bottom() (xfer.model_handler.ModelHandler method), 70
 add_layer_top() (xfer.model_handler.ModelHandler method), 70
 args (xfer.model_handler.exceptions.ModelArchitectureError attribute), 72
 args (xfer.model_handler.exceptions.ModelError attribute), 72

B

BATCHNORM (xfer.model_handler.consts.LayerType attribute), 72
 BnnRepurposer (class in xfer), 60

C

CONCAT (xfer.model_handler.consts.LayerType attribute), 72
 CONVOLUTION (xfer.model_handler.consts.LayerType attribute), 72

D

DATA (xfer.model_handler.consts.LayerType attribute), 73
 deserialize() (xfer.BnnRepurposer method), 62
 deserialize() (xfer.GpRepurposer method), 58
 deserialize() (xfer.LrRepurposer method), 54
 deserialize() (xfer.MetaModelRepurposer method), 50
 deserialize() (xfer.NeuralNetworkFineTuneRepurposer method), 64
 deserialize() (xfer.NeuralNetworkRandomFreezeRepurposer method), 66
 deserialize() (xfer.NeuralNetworkRepurposer method), 52
 deserialize() (xfer.Repurposer method), 48
 deserialize() (xfer.SvmRepurposer method), 56

drop_layer_bottom() (xfer.model_handler.ModelHandler method), 70
 drop_layer_top() (xfer.model_handler.ModelHandler method), 70
 DROPOUT (xfer.model_handler.consts.LayerType attribute), 72

E

EMBEDDING (xfer.model_handler.consts.LayerType attribute), 73
 end_annealing (xfer.BnnRepurposer attribute), 61

F

feature_layer_names (xfer.BnnRepurposer attribute), 61
 feature_layer_names (xfer.GpRepurposer attribute), 58
 feature_layer_names (xfer.LrRepurposer attribute), 54
 feature_layer_names (xfer.MetaModelRepurposer attribute), 49
 feature_layer_names (xfer.SvmRepurposer attribute), 56
 FLATTEN (xfer.model_handler.consts.LayerType attribute), 72
 FULLYCONNECTED (xfer.model_handler.consts.LayerType attribute), 72

G

get_features_from_source_model() (xfer.BnnRepurposer method), 62
 get_features_from_source_model() (xfer.GpRepurposer method), 58
 get_features_from_source_model() (xfer.LrRepurposer method), 54
 get_features_from_source_model() (xfer.MetaModelRepurposer method), 50
 get_features_from_source_model() (xfer.SvmRepurposer method), 56
 get_layer_names_matching_type() (xfer.model_handler.ModelHandler method), 71
 get_layer_output() (xfer.model_handler.ModelHandler method), 71

[get_layer_parameters\(\)](#) (xfer.model_handler.ModelHandler method), [71](#)
[get_layer_type\(\)](#) (xfer.model_handler.ModelHandler method), [70](#)
[get_module\(\)](#) (xfer.model_handler.ModelHandler method), [70](#)
[get_params\(\)](#) (xfer.BnnRepurposer method), [61](#)
[get_params\(\)](#) (xfer.GpRepurposer method), [58](#)
[get_params\(\)](#) (xfer.LrRepurposer method), [54](#)
[get_params\(\)](#) (xfer.MetaModelRepurposer method), [49](#)
[get_params\(\)](#) (xfer.NeuralNetworkFineTuneRepurposer method), [64](#)
[get_params\(\)](#) (xfer.NeuralNetworkRandomFreezeRepurposer method), [66](#)
[get_params\(\)](#) (xfer.NeuralNetworkRepurposer method), [51](#)
[get_params\(\)](#) (xfer.Repurposer method), [48](#)
[get_params\(\)](#) (xfer.SvmRepurposer method), [56](#)
[GpRepurposer](#) (class in xfer), [57](#)

L

[layer_names](#) (xfer.model_handler.ModelHandler attribute), [71](#)
[LayerType](#) (class in xfer.model_handler.consts), [72](#)
[LRN](#) (xfer.model_handler.consts.LayerType attribute), [73](#)
[LrRepurposer](#) (class in xfer), [53](#)

M

[MEAN](#) (xfer.model_handler.consts.LayerType attribute), [73](#)
[MetaModelRepurposer](#) (class in xfer), [48](#)
[ModelArchitectureError](#), [72](#)
[ModelError](#), [72](#)
[ModelHandler](#) (class in xfer.model_handler), [69](#)

N

[NeuralNetworkFineTuneRepurposer](#) (class in xfer), [63](#)
[NeuralNetworkRandomFreezeRepurposer](#) (class in xfer), [65](#)
[NeuralNetworkRepurposer](#) (class in xfer), [50](#)

P

[PLUS](#) (xfer.model_handler.consts.LayerType attribute), [72](#)
[POOLING](#) (xfer.model_handler.consts.LayerType attribute), [72](#)
[predict_label\(\)](#) (xfer.BnnRepurposer method), [62](#)
[predict_label\(\)](#) (xfer.GpRepurposer method), [59](#)
[predict_label\(\)](#) (xfer.LrRepurposer method), [54](#)
[predict_label\(\)](#) (xfer.MetaModelRepurposer method), [50](#)
[predict_label\(\)](#) (xfer.NeuralNetworkFineTuneRepurposer method), [64](#)
[predict_label\(\)](#) (xfer.NeuralNetworkRandomFreezeRepurposer method), [66](#)
[predict_label\(\)](#) (xfer.NeuralNetworkRepurposer method), [52](#)
[predict_label\(\)](#) (xfer.Repurposer method), [48](#)
[predict_label\(\)](#) (xfer.SvmRepurposer method), [56](#)
[predict_probability\(\)](#) (xfer.BnnRepurposer method), [62](#)
[predict_probability\(\)](#) (xfer.GpRepurposer method), [59](#)
[predict_probability\(\)](#) (xfer.LrRepurposer method), [54](#)
[predict_probability\(\)](#) (xfer.MetaModelRepurposer method), [49](#)
[predict_probability\(\)](#) (xfer.NeuralNetworkFineTuneRepurposer method), [64](#)
[predict_probability\(\)](#) (xfer.NeuralNetworkRandomFreezeRepurposer method), [66](#)
[predict_probability\(\)](#) (xfer.NeuralNetworkRepurposer method), [51](#)
[predict_probability\(\)](#) (xfer.Repurposer method), [48](#)
[predict_probability\(\)](#) (xfer.SvmRepurposer method), [57](#)

R

[repurpose\(\)](#) (xfer.BnnRepurposer method), [62](#)
[repurpose\(\)](#) (xfer.GpRepurposer method), [59](#)
[repurpose\(\)](#) (xfer.LrRepurposer method), [54](#)
[repurpose\(\)](#) (xfer.MetaModelRepurposer method), [49](#)
[repurpose\(\)](#) (xfer.NeuralNetworkFineTuneRepurposer method), [64](#)
[repurpose\(\)](#) (xfer.NeuralNetworkRandomFreezeRepurposer method), [66](#)
[repurpose\(\)](#) (xfer.NeuralNetworkRepurposer method), [51](#)
[repurpose\(\)](#) (xfer.Repurposer method), [48](#)
[repurpose\(\)](#) (xfer.SvmRepurposer method), [57](#)
[Repurposer](#) (class in xfer), [47](#)

S

[save_repurposer\(\)](#) (xfer.BnnRepurposer method), [62](#)
[save_repurposer\(\)](#) (xfer.GpRepurposer method), [59](#)
[save_repurposer\(\)](#) (xfer.LrRepurposer method), [54](#)
[save_repurposer\(\)](#) (xfer.MetaModelRepurposer method), [50](#)
[save_repurposer\(\)](#) (xfer.NeuralNetworkFineTuneRepurposer method), [64](#)
[save_repurposer\(\)](#) (xfer.NeuralNetworkRandomFreezeRepurposer method), [66](#)
[save_repurposer\(\)](#) (xfer.NeuralNetworkRepurposer method), [52](#)
[save_repurposer\(\)](#) (xfer.Repurposer method), [48](#)
[save_repurposer\(\)](#) (xfer.SvmRepurposer method), [57](#)
[save_symbol\(\)](#) (xfer.model_handler.ModelHandler method), [71](#)
[serialize\(\)](#) (xfer.BnnRepurposer method), [62](#)
[serialize\(\)](#) (xfer.GpRepurposer method), [58](#)
[serialize\(\)](#) (xfer.LrRepurposer method), [54](#)
[serialize\(\)](#) (xfer.MetaModelRepurposer method), [50](#)
[serialize\(\)](#) (xfer.NeuralNetworkFineTuneRepurposer method), [65](#)

[serialize\(\) \(xfer.NeuralNetworkRandomFreezeRepurposer method\)](#), [67](#)
[serialize\(\) \(xfer.NeuralNetworkRepurposer method\)](#), [52](#)
[serialize\(\) \(xfer.Repurposer method\)](#), [48](#)
[serialize\(\) \(xfer.SvmRepurposer method\)](#), [56](#)
[SOFTMAXOUTPUT \(xfer.model_handler.consts.LayerType attribute\)](#), [72](#)
[source_model \(xfer.BnnRepurposer attribute\)](#), [62](#)
[source_model \(xfer.GpRepurposer attribute\)](#), [59](#)
[source_model \(xfer.LrRepurposer attribute\)](#), [55](#)
[source_model \(xfer.MetaModelRepurposer attribute\)](#), [49](#)
[source_model \(xfer.SvmRepurposer attribute\)](#), [57](#)
[start_annealing \(xfer.BnnRepurposer attribute\)](#), [61](#)
[SVMOUTPUT \(xfer.model_handler.consts.LayerType attribute\)](#), [73](#)
[SvmRepurposer \(class in xfer\)](#), [55](#)

U

[update_sym\(\) \(xfer.model_handler.ModelHandler method\)](#), [71](#)

V

[visualize_net\(\) \(xfer.model_handler.ModelHandler method\)](#), [71](#)

W

[with_traceback\(\) \(xfer.model_handler.exceptions.ModelArchitectureError method\)](#), [72](#)
[with_traceback\(\) \(xfer.model_handler.exceptions.ModelError method\)](#), [72](#)

X

[xfer.model_handler.consts \(module\)](#), [72](#)
[xfer.model_handler.exceptions \(module\)](#), [71](#)