Practical 4

Jumping Rivers

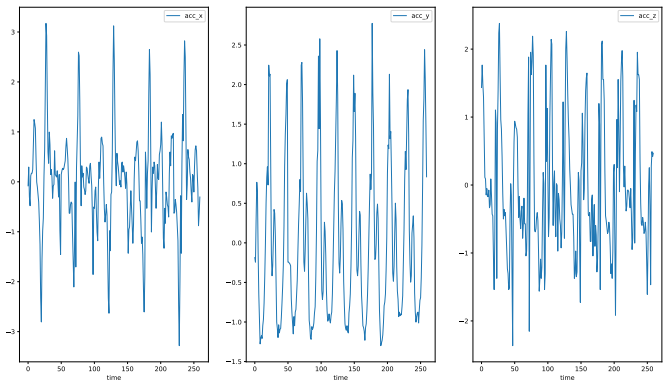
Predict a person based on accelerometer data from walking

We have accelerometer data on 15 individuals who all walked for a period of time. Each observation is a set of values from the x,y and z axis of an accelerometer in 5 seconds sampled at a frequency of 52Hz. One sample is 260 observations. We have between 300-600 observations on each individual

```
import jrpytensorflow

walking = jrpytensorflow.datasets.load_walking()
```

The following code will produce a visualisation of one sample

```
import matplotlib.pyplot as plt
import numpy as np

sub = walking[walking['sample'] == 400]
sub.loc[:,'time'] = np.arange(260)

fig, (ax1,ax2,ax3) = plt.subplots(1,3, figsize = (18,10))
sub.plot(x = 'time', y = 'acc_x', ax = ax1)
sub.plot(x = 'time', y = 'acc_y', ax = ax2)
sub.plot(x = 'time', y = 'acc_z', ax = ax3)
plt.show()
```



At present this data is not in a particularly convenient structure for training my model. The following code will create the (n,260,3) (n observations of 260 inputs for 3 channels) input shape and class labels as seperate array objects

```
dims = ['acc_x','acc_y','acc_z']
x = np.dstack([walking[[d]].values.reshape(-1,260) for d in dims])
y = walking['person'].values[::260] - 1
```

Each observation is a sequence of 260 values with 3 channels. This is the sort of data structure where we would use a convolutional neural network with 1d convolutions.

- Create a model structure that takes in the 260 input features for each of 3 data channels and returns 15 output features (one for each person). We will want a set of convolution and pooling layers to begin with before having some linear layers to get to the final output. The convolutions and pooling extract features from the 3 channels of sequences, the linear layers then map those features to the final output.
- partition your data into training and test sets and binarize the labels (Optional: partition data into a validation set as well)
- compile and train your model
- Using the `.history` object, can you plot how the loss and accuracy changes over the training time
- what is the loss and accuracy in your test set?
- Try to view the model architecture in TensorBoard

For me this gives 95% plus accuracy on classifying a person purely on the accelerometer data while walking. Pretty neat!

- In the image example in the notes we were able to use `ImageDataGenerator()` to augment the data and fit it using `fit_generator`, allowing us to pass the data in pieces rather than store the dataset entirely in memory. Keras also has `TimeseriesGenerator()` - a generator for times series data. Using this and `.fit_generator` can you fit your model?