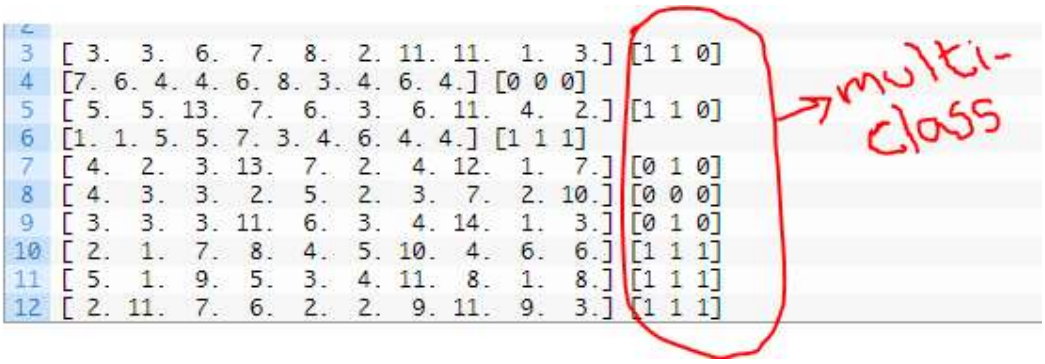
Multi-label classification using Deep Neural Networks

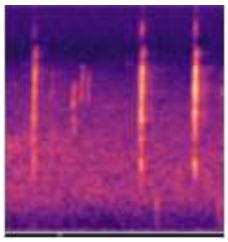
We are converting our flac audio files to melspectogram for better analysis. These are 5 second clips and our current system is detecting the most prominent animal sound in these clips using single label classification via CNN. However, we might want to detect overlapping and background sounds from other animals, for which we need multi-label classification models. An efficient way of achieving this may be using multi label deep neural networks (ML DNN) as showcased in the IEEE paper below:

https://ieeexplore-ieee-org.ezproxy-b.deakin.edu.au/stamp/stamp.jsp?tp=&arnumber=7362845

They have compared using multi-layer DNN and combined single layer DNN. While the later is slightly less in terms of accuracy, it offers more flexibility.



Using deep learning neural networks, we can predict the multiple classes in a single 5 second window as shown above. For example, a 5 second audio clip has sounds of a wild boar, an elephant and a bird overlapping at the same time:



Then the output label of this clip would read [1 1 1]. If there was only a wild boar, then it would read [1 0 0] and so on.

The deep neural networks for multi label classification will have one input layer with multiple inputs, 2 or more hidden layers and an output layer with multiple classes. Multi-label classification can be supported directly by neural networks simply by specifying the number of target labels there is in the problem as the number of nodes in the output layer. For example, a task that has three output labels (classes) will require a neural network output layer with three nodes in the output layer.

To configure a neural network model for multi-label classification, the requirements are:

1. Number of nodes in the output layer matches the number of labels.
2. Sigmoid activation for each node in the output layer.
3. Binary cross-entropy loss function.

Additionally, it is good practice to use k-fold cross-validation instead of train/test splits of a dataset to get an unbiased estimate of model performance when making predictions on new data.

Information related to the definition and implementation of these DNNs can be found in the link below:

https://machinelearningmastery.com/multi-label-classification-with-deep-learning/