3.8 Assignment Solution

The size of the encoding layer should be less than 8 and more than 0 as we need to have a bottleneck from the previous layer (which was size 8). We have found: `EMBEDDING_SIZE = 2` seems to work the best! The encoding layer size is a critical factor in how much information our model can encode. Remember, we want the encoding layer (and the model in general) to be large enough that it can encode the normal signals that represent 90% of our training data but we want it to be small enough so that it can't learn to reproduce the 10% of our training data that is anomalous. For this dataset, the model and encoding layer can be very small, but it's important to remember that this is a very controlled and simple example.

The threshold depends on the training and on the encoding layer size but somewhere in the range of `threshold = 0.037` works best. When picking the threshold the goal is to maximize the accuracy, precision, and recall. The 'knee' of the ROC curve is a good place to start as it represents a good balance between precision and recall. For this application, it might be more important to properly classify an abnormal rhythm as such, instead of maximizing the accuracy.

As always if you would like to go back and try the assignment again, it can be found here:
https://colab.research.google.com/github/tinyMLx/colabs/blob/master/3-8-16-Assignment.ipynb