Part 3: heart

Test/Train Split
I split the last 1/7 of the data off for testing and used the rest as training. This is similar to the ratios on previous datasets and feels reasonable.

Category Data
I pre-processed it by changing the 'famhist' (family history) to 0 or 1 depending if it was 'Absent' or 'Present'. This is done directly in the numpy DataFrame. The tutorial provided in the assignment was too complex. While I see the advantage of having the categorical data processing baked into the model itself, I was not able to implement it.

Overfitting
Originally, after 30 epochs, the model had 0.9773 accuracy on the training set but 0.7576 on the testing set. This is evidence of overfitting.

Reduce Overfitting
I added a L2 regularizer on the first and second Dense layers to reduce overfitting. This helps because, by adding a cost to large weights, it discourages to model from becoming over-reliant on one variable. I have a larger penalty on the first Dense layer because the I believe that it should heavily avoid high weights on the raw data, while the second layer is recieving information that is already processed and might be able to use slightly larger weights.

Results
The new result is 0.8561 accuracy on training set and 0.8030 on testing set after the same 15 epochs, which is much less overfit.
Layers: normalizer, Dense 128, Dense 64, Dense 1 sigmoid.
Dense 128 has L2 regularization with penalty constant 0.005, Dense 64 has L2 regularization with penalty constant 0.001.
15 epochs. Batch size 1 because the dataset is small and I can afford to run small batches.