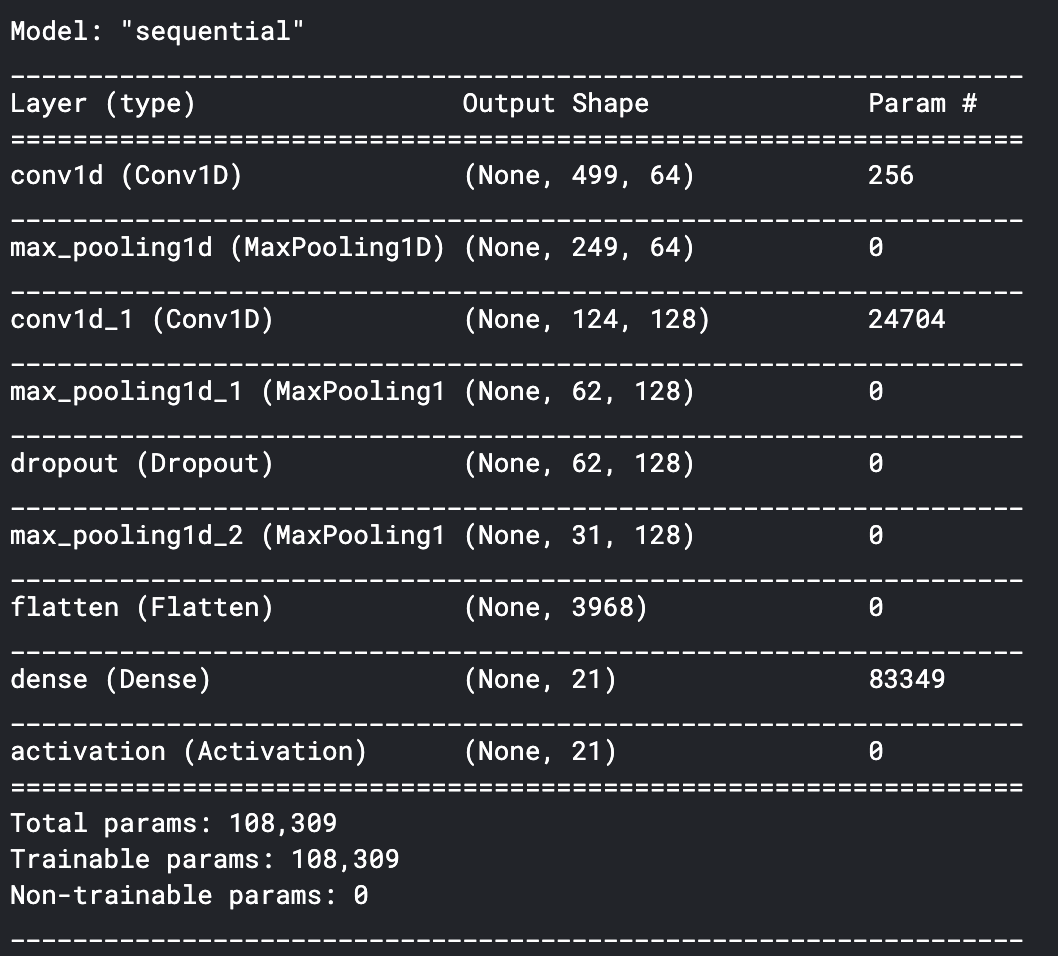
Model Architecture:



Metrics:

* Accuracy: 0.8210796915167096
* Balanced accuracy: 0.7245093590667615
* Precision: 0.8364227943206033
* Recall: 0.8210796915167096
* F1: 0.8103630068659465

Convolutional Neural Networks are a family of deep learning algorithms geared towards classifying two-dimensional data. They are feedforward neural networks which means that as new data is fed and used to train the model, previous vectors and features are forgotten. CNN’s, compared to artificial neural networks made of fully connected layers, keep their structure(edges and gradients) local. Within each convolutional layer, hence their name, the model applies a convolution to each tile in the data at once with the same kernel. Each layer is able to obtain more complex data, such as patterns and shapes, which is then used for predictions. Although CNN’s are mainly geared towards image data, we trained them on our one-dimensional opcode data and discovered that they performed well. For our 1D CNN model, we chose to have 2 Convolutional layers(the first layer having 64 filters followed by a second layer with 128 filters). Both layers have a kernel size of 3 and a stride of 2. Following each convolutional layer, we added Max Pooling layers with a kernel size of 2 to reduce feature dimensionality and reduce overfitting. We then integrated a Dropout layer of 90%, where we intentionally ignore neurons to further prevent overfitting and follow this layer with another Max Pooling layer. Finally, we add a fully connected Dense layer of 21, the number of families we are classifying, to make predictions. While training our model, we used a batch size of 32 and around 100 epochs.