# Maria Musiał 156062 – Analysis of a model

Model is designed to classify dog and cats images.

It uses a feedforward neural network with 3 hidden layers.

Input layer: image vector

3 fully connected hidden layers that use ReLU as activation function

Output layer with classification

First initialized weights are using standard deviation, so they are sensible.

Images at first are flattened to 2D array (of length width\*height\*#channels), that has a single image as a row, number of rows being number of images.

Then as a vector they are multiplied by initialized weights, add bias and perform Relu.

This is performed for 3 layers 1024, 256, 64 neurons each.

Last layer- output layer has softmax function, as its used for probabilities summing to 1 (case of classification)

Role of the training rule:

We use sparse softmax cross entropy as a loss function to match how well is the model performing in epochs.

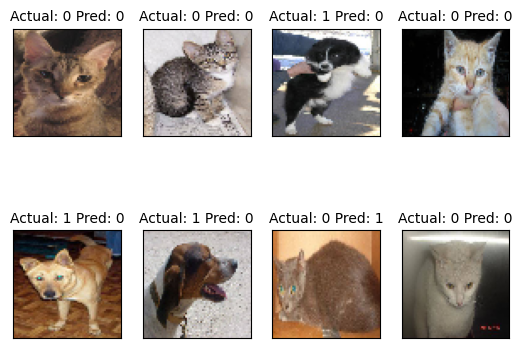
Adam optimizer with LR=0.001 , beta1= 0.9 and beta2=0.999 is used

Higher learning rate – larger steps during each iteration, potentially speeding up the training, but risking overshooting minima. Lower LR can be slower, but more accurate (smaller steps)

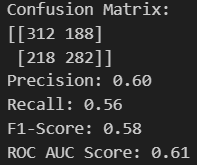
Results:

Training duration: 38min.

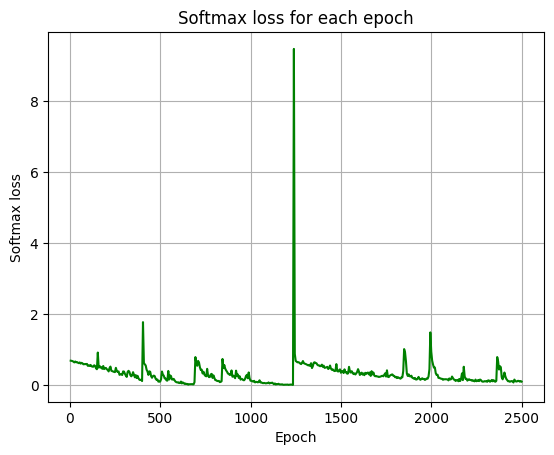
Some classified examples from test set:

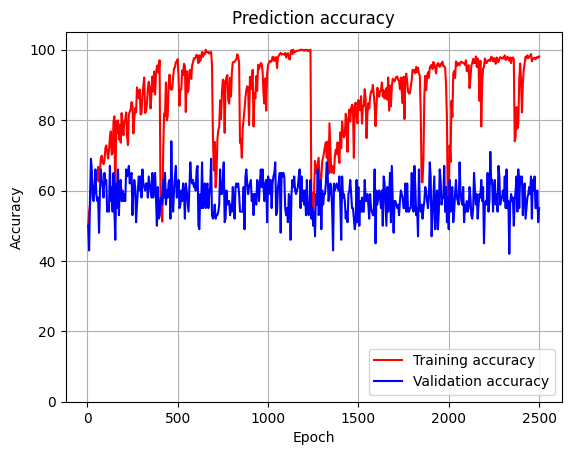


Data checked on whole test set:



Validation curves:





We can see that randomly indexing data that has high variety gives us fluctuations in the plots.

From the graphs we can conclude overfitting or complexe data structues, as we have drops. I’m betting on complexe data, as those are images with features to extract, which isn’t done here.