

Data-X – Homework 10

Louis TILLOY

Problem 1

Final test accuracy : 97.32%
Final test loss : 0.0874

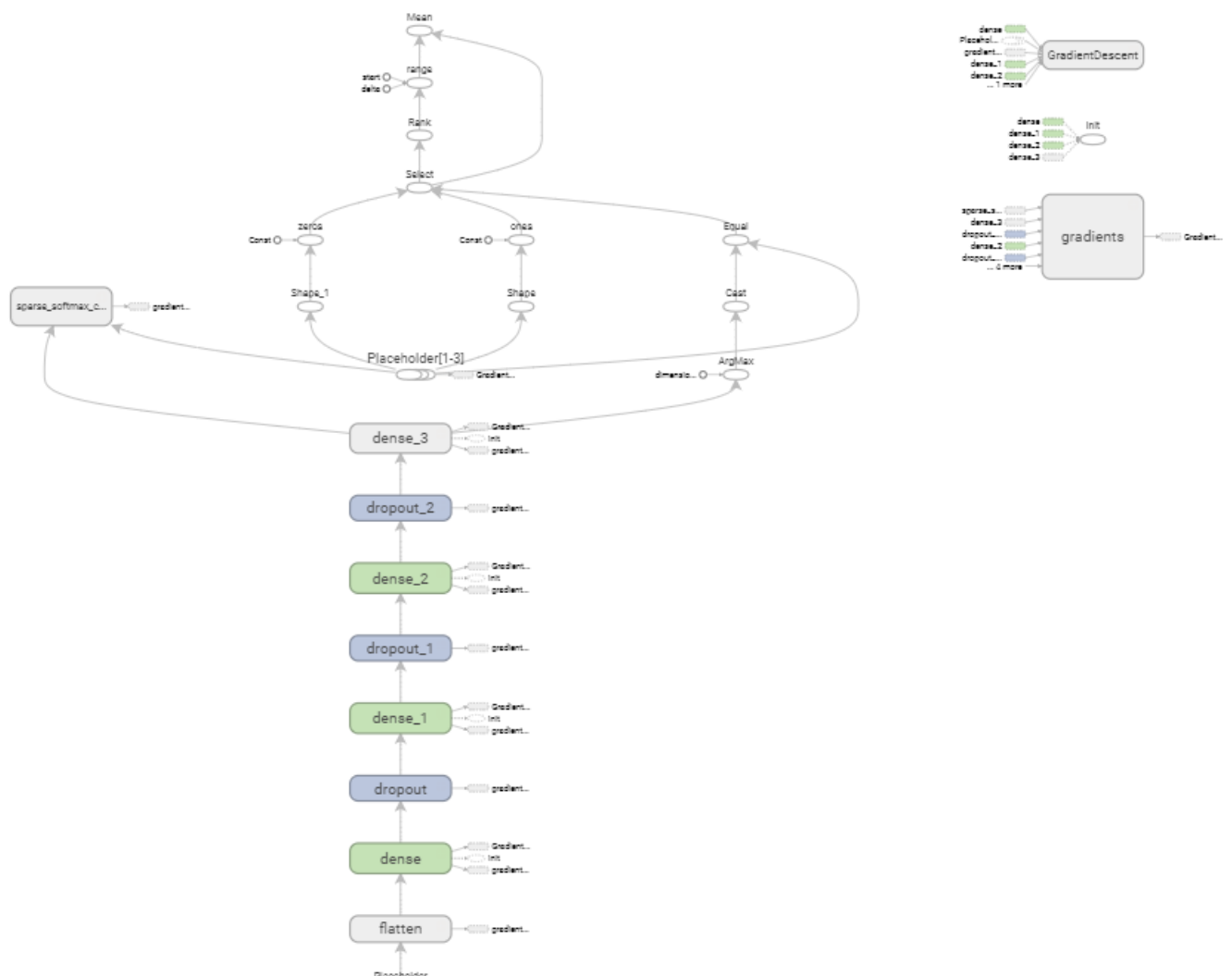


FIGURE 1 – Tensorboard graph of the neural network

Problem 2

For this problem, I used 2 convolutional layers with 64 and 32 filters. Each layer has 5×5 kernel size, "same" padding, batch normalization, max pooling of strides 2/kernel 2×2 and a 0.7 dropout rate. I used the function swish for the activation function (see figure 3).

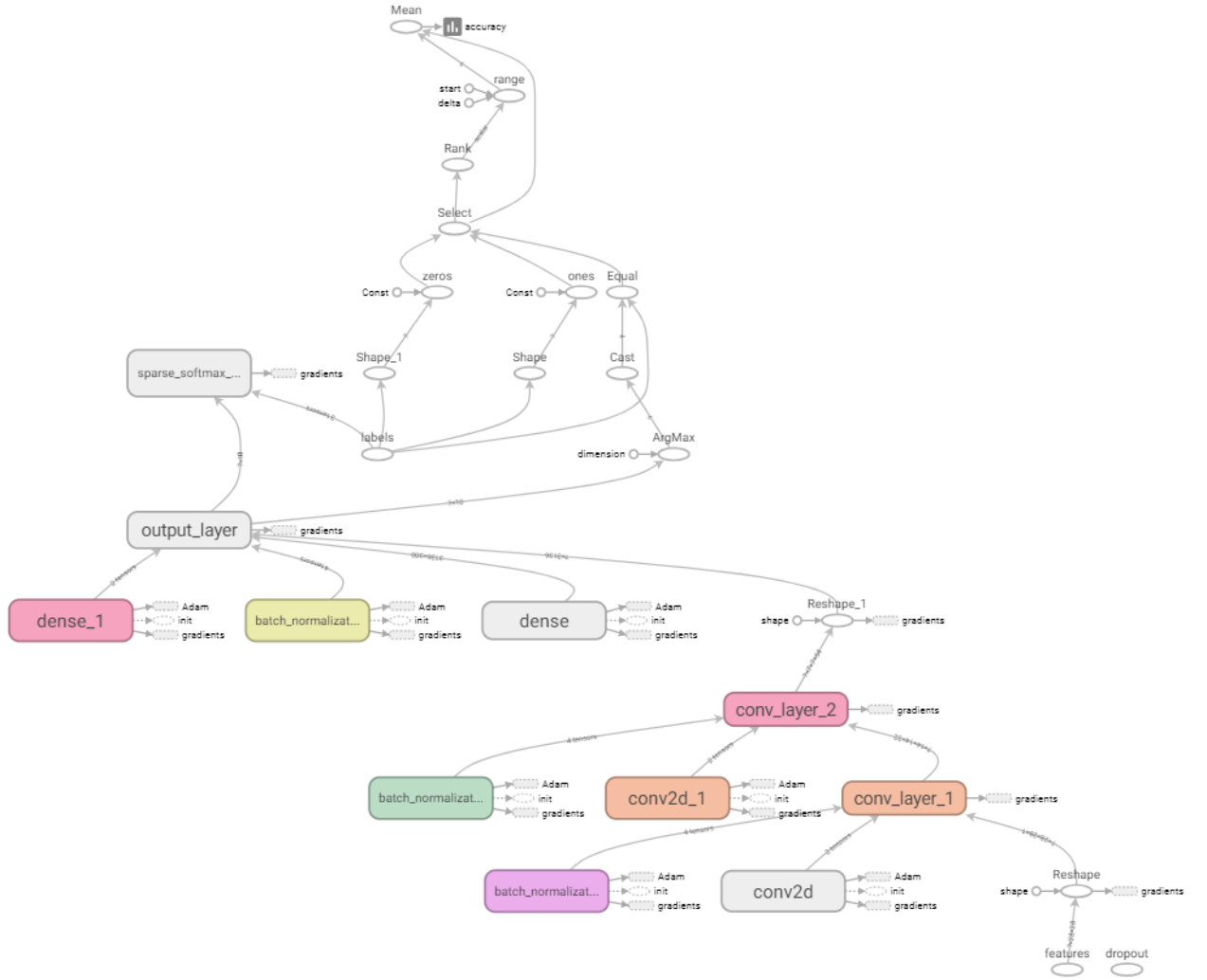


FIGURE 2 – Tensorboard graph of the neural network

I also used 2 fully connected layers of 300 units, still with batch normalization and dropout equal to 0.7 between the 2 layers. I used random uniform initialization between -0.1 and 0.1 to initialize the weights of these fully connected layers.

For the training phase, I used a batch size of 32, an Adam Optimizer with the default hyper-parameters and trained the network for 40 epochs.

Final test accuracy : 99.25%
Final test loss : 0.07003816

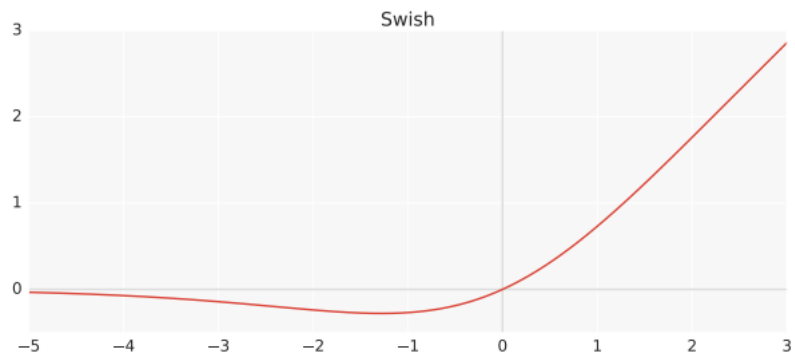


FIGURE 3 – Swish function

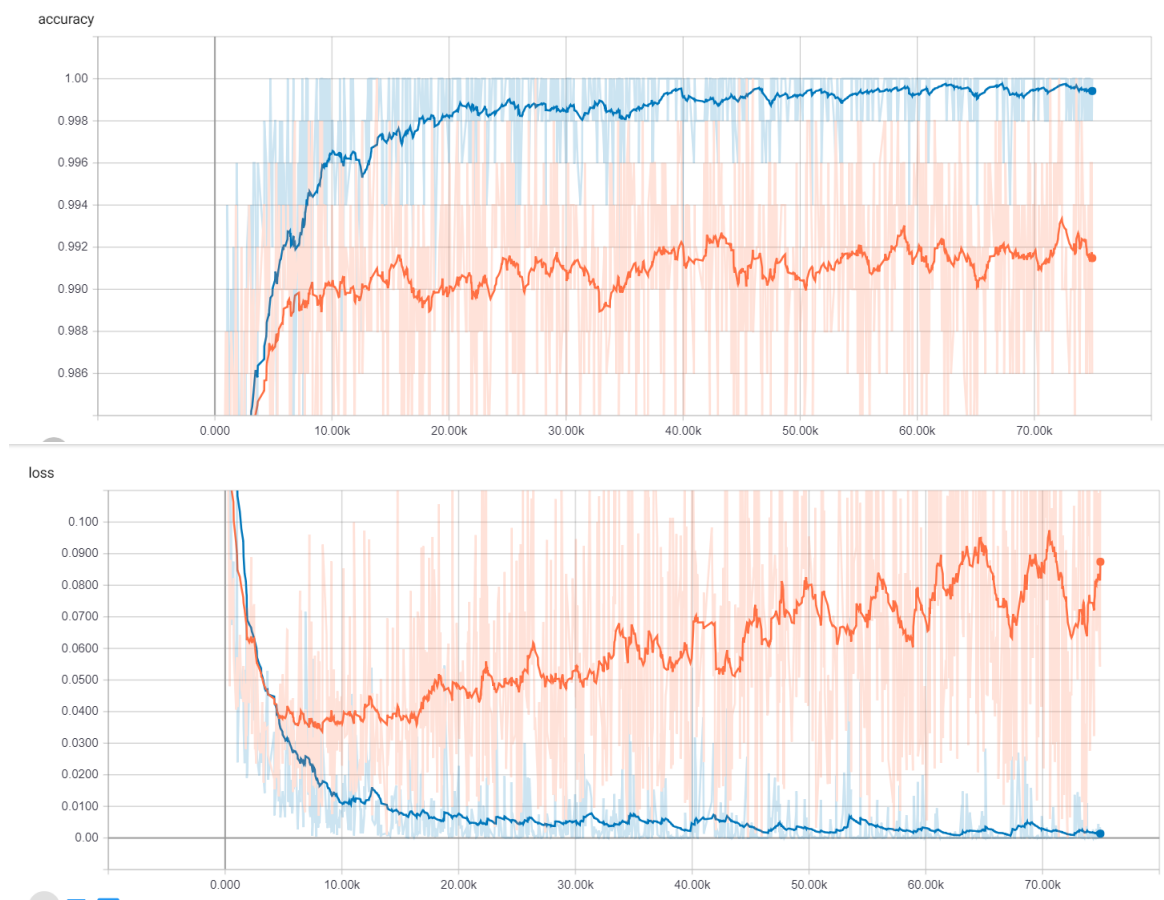


FIGURE 4 – Evolution of the accuracy and loss on the training set and the testing set in training in Tensorboard