Convolutional Neural Network

# Implement a convolutional “neural” network consisting of a sequence of convolution and pooling transforms. The network output must end with a SoftArgMax transformation. Try to find the best network architecture. To minimize, use one of the adaptive gradient descent methods.

# Cross Entropy should be used as the error function to be minimized, and Error Rate as the control error function. Also build a confusion matrix.

# Datasets

Use the [MNIST](http://yann.lecun.com/exdb/mnist/) dataset to find the best architecture. Train and test the best found network architecture on the [Fashion-MNIST](https://github.com/zalandoresearch/fashion-mnist) dataset