Project Proposal

MINIST Dataset

The MINIST dataset has a training dataset and a test dataset, with 60,000 and 10,000 samples, respectively. we split the training dataset into a training dataset of 50,000 samples and a validation dataset of 10,000. We train on the training dataset, stop training via the validation dataset, and test on the test dataset.

Methods

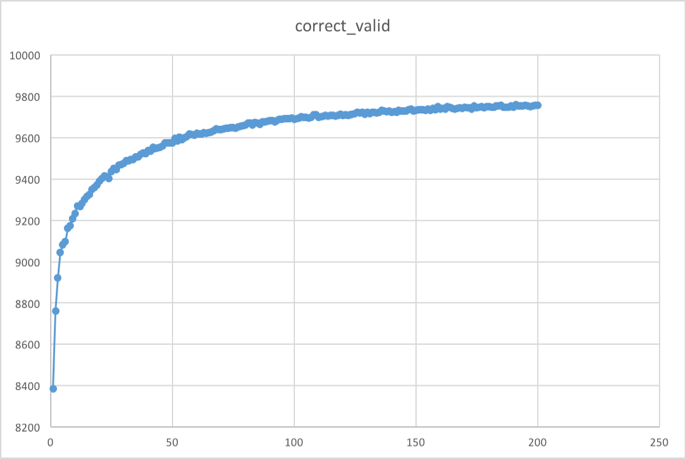
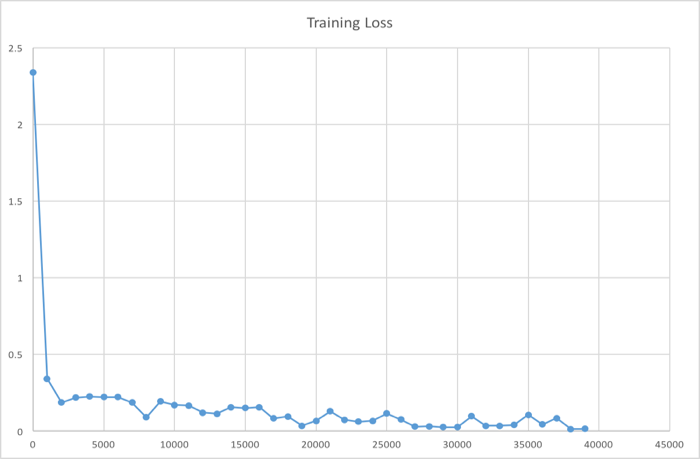
We use two methods to do the classification task on the dataset. One is an end-to-end model, which is a 2-hidden Feedforward network. The input of the network is a normalized vector of 28\*28 dimensions, with values in [0,1]. The nodes of the two hidden layer are both 512. The activation of the last layer is the softmax function. And the loss function of the network is cross-entropy. The classification accuracy of test dataset is directly got by the trained network. The other one is not an end-to-end model, which firstly extract the features of images with the trained 2-hidden Feedforward network above, secondly train a SVM classifier with the extracted features. For the test dataset, we classify images by the trained SVM classifier with the features extracted by the trained network.

Group members:

郭倩(网数)

段雪野(网数)

Project Report

 The end-to-end model has two problems: (1) parameters need to be update rapidly (2) a careful initialization is needed to converge with less updating. Since the size of the parameters of the network are 785\*512+513\*512+512\*10 and the training dataset is large, updating with gradient descent is inefficient. So it is same with the stochastic gradient descent, for the variance between different training samples is large. So we adopt batch gradient descent with the size of the batch of 100. To solve the second problem, we initialize the parameters of each layer with a uniform distribution of , where is the number of nodes in the i-th layer. With the settings above, the training converges until 30,000–th updating, and the validation accuracy is 97.57%, test accuracy is 97.23%.

For the not end-to-end model, we train the nonlinear SVM classifier with 4-folds cross-validation on the validation set. Furthermore, the nonlinear function is RBF, which sets gamma= 0.0051, c=1. The accuracy of train set and test set are 97.48% and 97.2%, respectively (best result compared to the second best gamma= 0.0026 and c=1 with cross validation accuracy 97.28% and test accuracy 97.02%).