cs577 Assignment 2: Report

Yuanxing Cheng, A20453410, CS577-f22
Department of Mathematics
Illinois Institute of Technology

October 27, 2022

Problem statement

mnist 10 class classification and spam data binary classification problem. And tune hyperparameters to get better result.

Proposed solution

using hyperas and hyperopt to tune hyperparameters.

Implementation details

For mnist dataset, the validation dataset use the 5000 first training data. For spam dataset, the validation dataset use the first 400 entries of the training dataset obtained using *train test split* from scikie-learn.

The posssible hyperparameters are:

- second layer shape: 256, 512, 1024
- second layer activation: relu, sigmoid
- regularization: for both first and second layer, add dropout layer, batch normalizing layer, or add weight decay to second layer

Results and discussion

For mnist dataset, the best model uses the following architecture.

Layer type	output shape	parameter number
Dense	512	401920
batch normalizing	512	2048
active	512	0
dense	512	262656
active	512	0
batch normalizing	512	2048
dense	10	5130
active	10	0

The best model has accuracy 0.9551.
for spam dataset, the best model uses the following architecture.

Layer type	output shape	parameter number
Dense	512	29696
active	512	0
dense	1024	525312
active	1024	0
dense	1	1025
active	1	0

The best model has accuracy 0.9261.