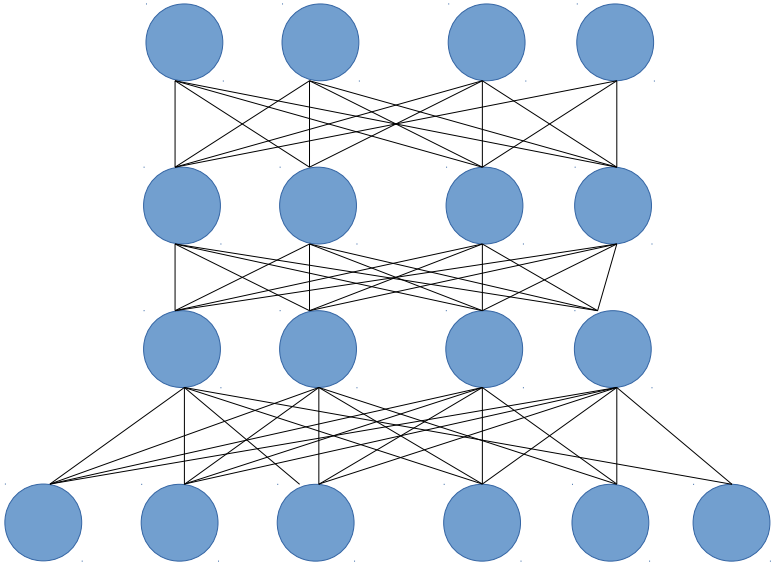
빅바이오 : part3-4 회차
(딥러닝 실습) TensorFlow 을 사용해서 암 환자 유전체에서 암 예측 모델 개발 (MLP)

한성국 2016/ 5/ 13

Multilayer Perceptron: Multinomial classification



- Input layer: 20502 neurons
- Hidden layer 1: 128 neurons
- Hidden layer 2: 128 neurons
- ...
- Multinomial classification: Softmax Layer: 34 neurons

Dataset

- Input features: 20502 (mRNA expression values)
- Train set: 7945
- Validation set: 1679
- Test set: 1679
- mRNA expression level: Normalization based on the number of input feature.

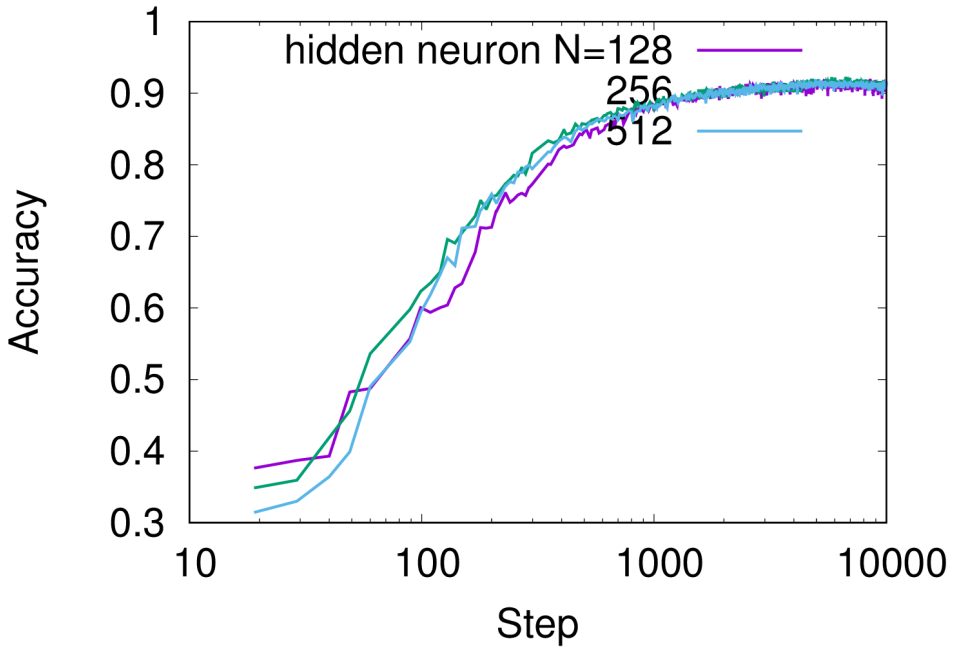
Hyper parameters and Ops

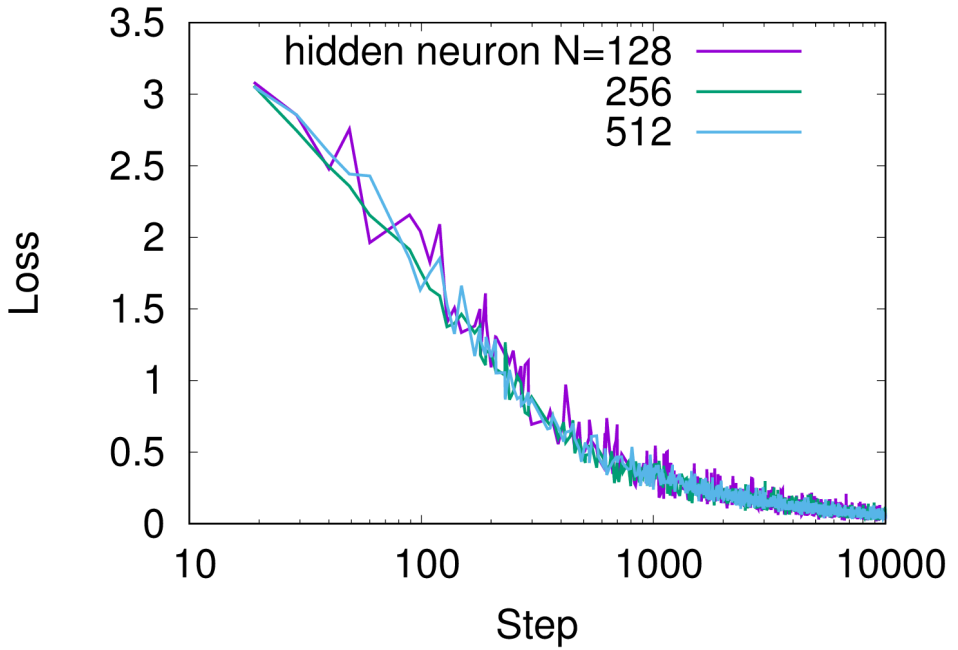
- Learning rate: 0.01
- Batch size 128
- Neuron activation function: RELU
- Loss function: cross entropy (softmax layer)
- Optimization: Stochastic Gradients Descent

Loss functions in TensorFlow

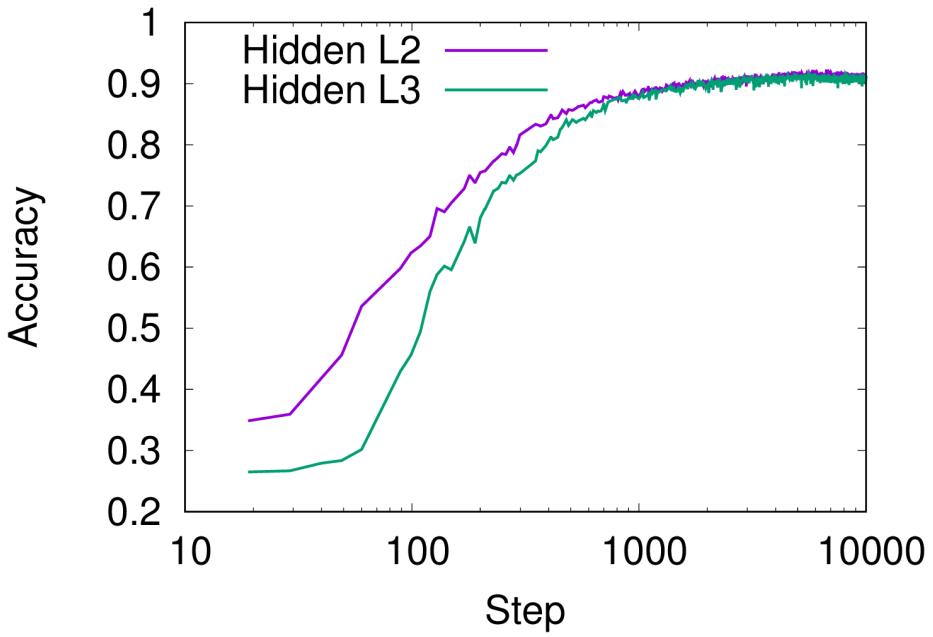
- When soft classes, use tf.nn.softmax_cross_entropy_with_logits(logits, labels, name=None)
- When classes are mutually exclusive, use tf.nn.sparse_softmax_cross_entropy_with_logits(logits, labels, name=None)

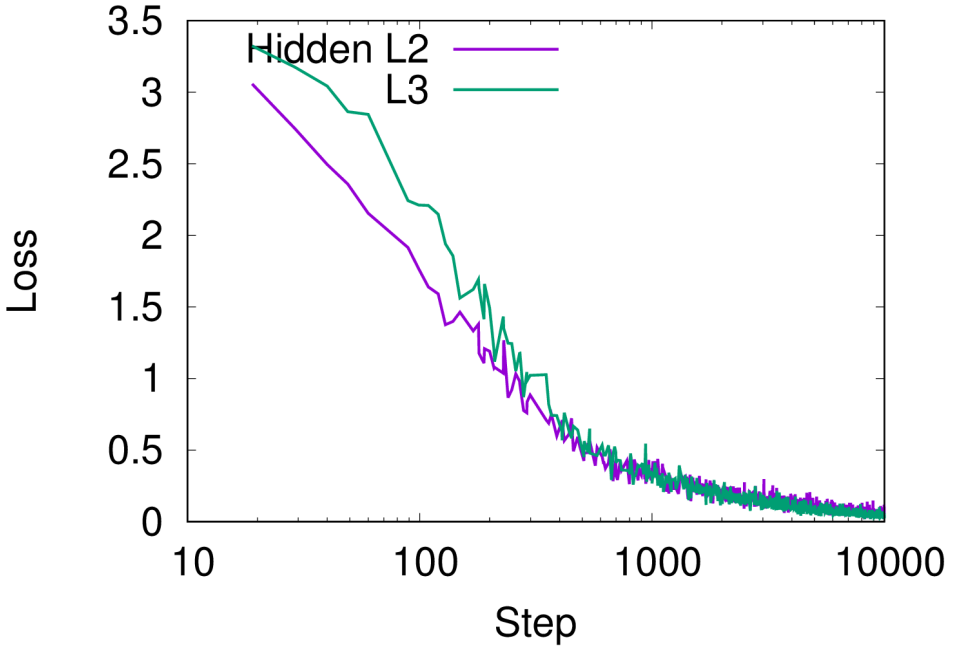
Training: loss and accuracy





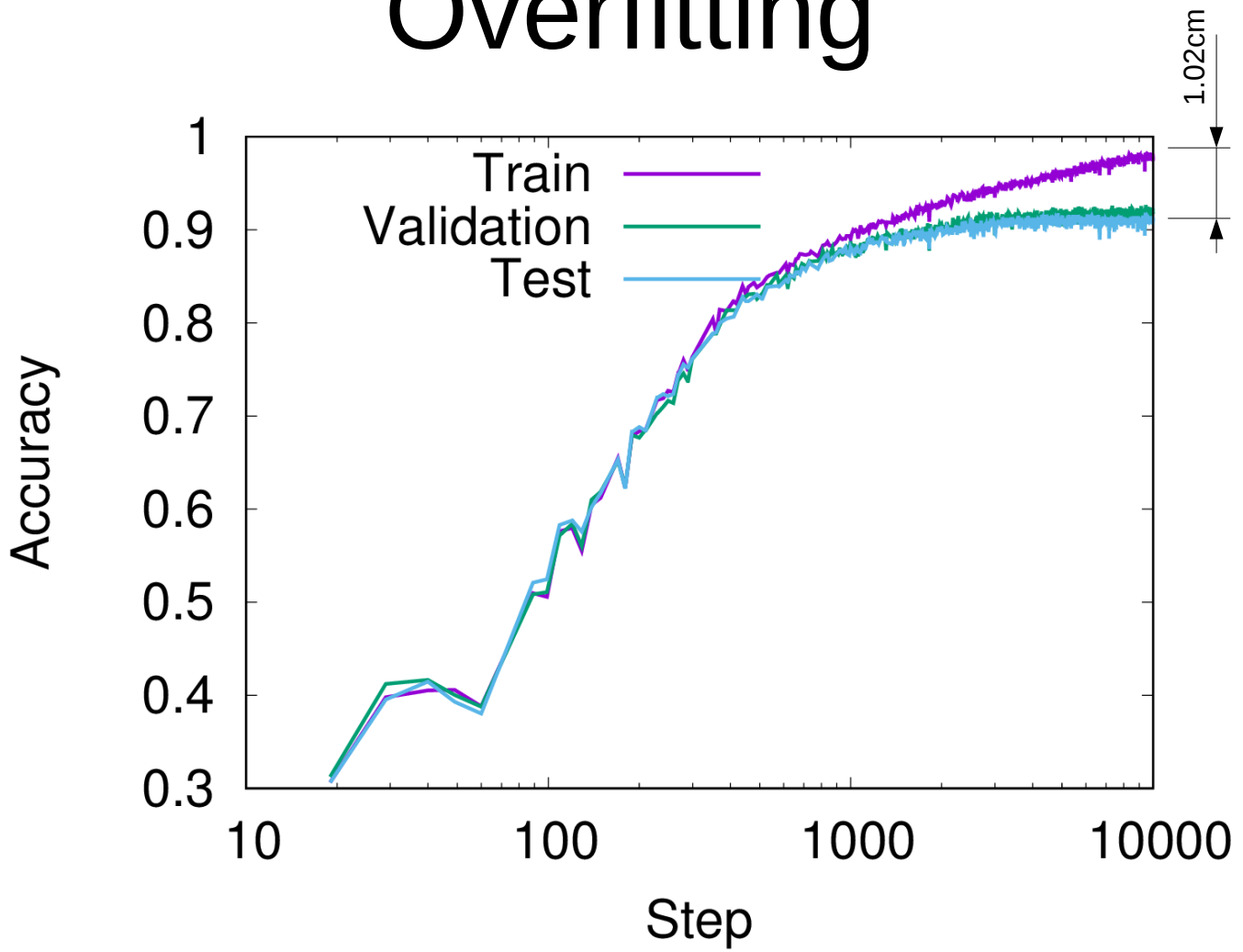
Hidden layer goes increases: Test accuracy





- Number of neurons in hidden layers: 256
- Batch sizes 128

Overfitting



- Hidden neurons=256
- Number of hidden Layers=3

Some more works...

- Batch size increases
- Use the different Optimizer: Adam, Adagrad,…
- How about to use CNN after change the input data to two dimensional.??