Soynet Week 4 2022312686

The Rise of Deep Learning Deep-Learning & Backpropagation

MLP & Backpropagation

• Perceptron(Weight, Bias) ➔ AND, OR (Solved) XOR (?) • Marvin Minksy ➔ Not by Single Perceptron, Maybe Multilayer? • But how to train MLP? (1st crisis)

• Solution ➔ Backpropagation (역전사) & Convolutinal Neural Networks (분할하여 인식)

• Scalability Issue (A Big Problem) → Backpropagation didn’t work well with lower layers (2nd crisis)

• Weight initialization in a clever way

Deep Learning

• Neural Network ➔ Deep Learning • ImageNet Classification

• Explain Photo

• Deep API Learning

• Speech recognition

• Etc

Lab 09-00 Tensor manipulation

Neural Network Introduction

Back Propagation

• �� = ���� + �� ➔����

����= ��,����

����= ��,����

����= 1

• Chain Rule

• �� = ���� + ��,����

����=����

����∗����

���� Utilize previous ‘local’ value.

• Tensor Flow use Chain Rule for Back Propagation. • Follow the Multiple Tensors …

•����������

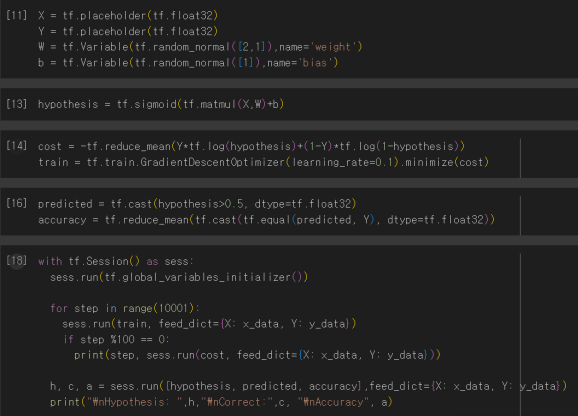
����0=����������

������×������−1

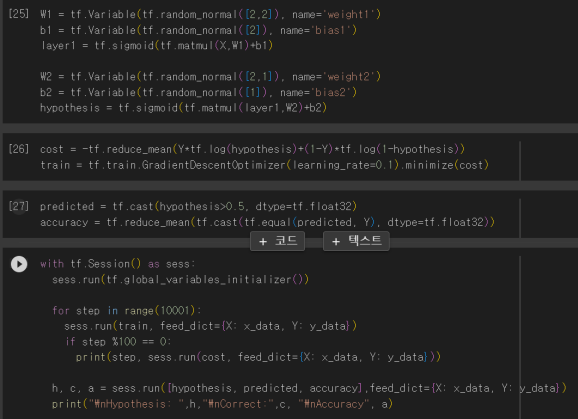
������−2× ⋯ ×����1

����0

Lab 09-01 Neural Net for XOR

Single 

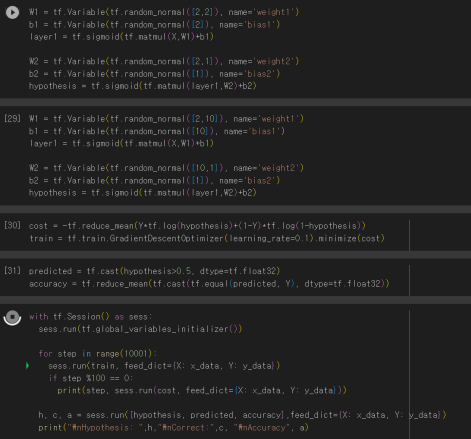
Layered NN Rate ➔ 0.5 …

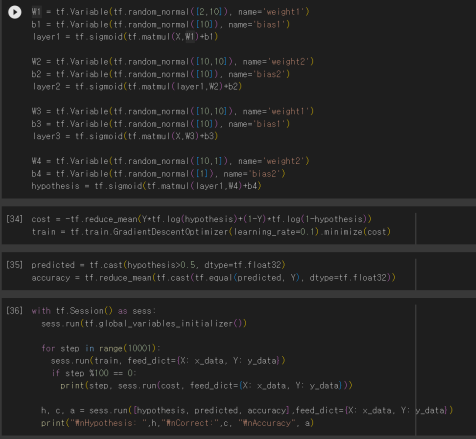
Multi 

Layered NN

Rate ➔ 1.0 !

Careful with Input Size

Wide Multi Layered NN

Wide Multi Layered NN

Lab 09-02 TensorBoard

How to Log

1. From the TF graph, decide which tensors you want to log 2. Merge all summaries

3. Create writer and add graph

4. Run summary merge and add\_summary

5. Launch TensorBoard

Exercise MNST

Neural Network ReLU

Activation Function; ReLU Function

• Drawback of Deep Learning

• so deep to be poor

• Sigmoid value is less than 1 ➔ with chain rule, final feedback would be lesser (Vanishing gradient)

• New function?

• ReLU (Rectified Linear Unit) Function

• �������� �� = max(0, ��)

• W/O final layer ➔ apply sigmoid

Activation Function; Variations

• Sigmoid Family

• tanh

• ReLU Family

• Leaky ReLU

• Maxout

• ELU

Neural Network Weight Initialization

Weight Initialization

• Why Important

• used for not only forward but also backward propagation • Initialize to zero result in fail ➔ no feed back

• Many Ways of Pre-Training

• RBM (Restricted Boatman Machine)

• Initialize weight for encoding and decoding value to be same • Pretrain weights sequentially (Fine Tuning)

• Xavier Initialization & He’s Initialization

• Make sure the weight are ‘just right’ not too big and small

Neural Network Overfitting

Solutions; Regularization & Dropout

• Regularization

• Restrain weight in proper range

�� ��ො − ��2

• ��2 ���������������������������� = σ��=1

�� ��ෞ��2

• �������� = ��������′ + �� σ��=1

• Minimalize Cost means make weight close to zero ➔ restrain weight • Control with lambda

• Dropout

• Randomly dropout some nodes in a layer in “training” • Not in Executing

