Flavor tagging based on deep learning

Bingyang Zhang Instructor:Gang Li

CONTENTS

01

The relationship among AI,machine learning and deep learning

AI,Machine Learning AND Deep Learning

AI(Artificial Intelligence)

AI is intelligence exhibited by machines.

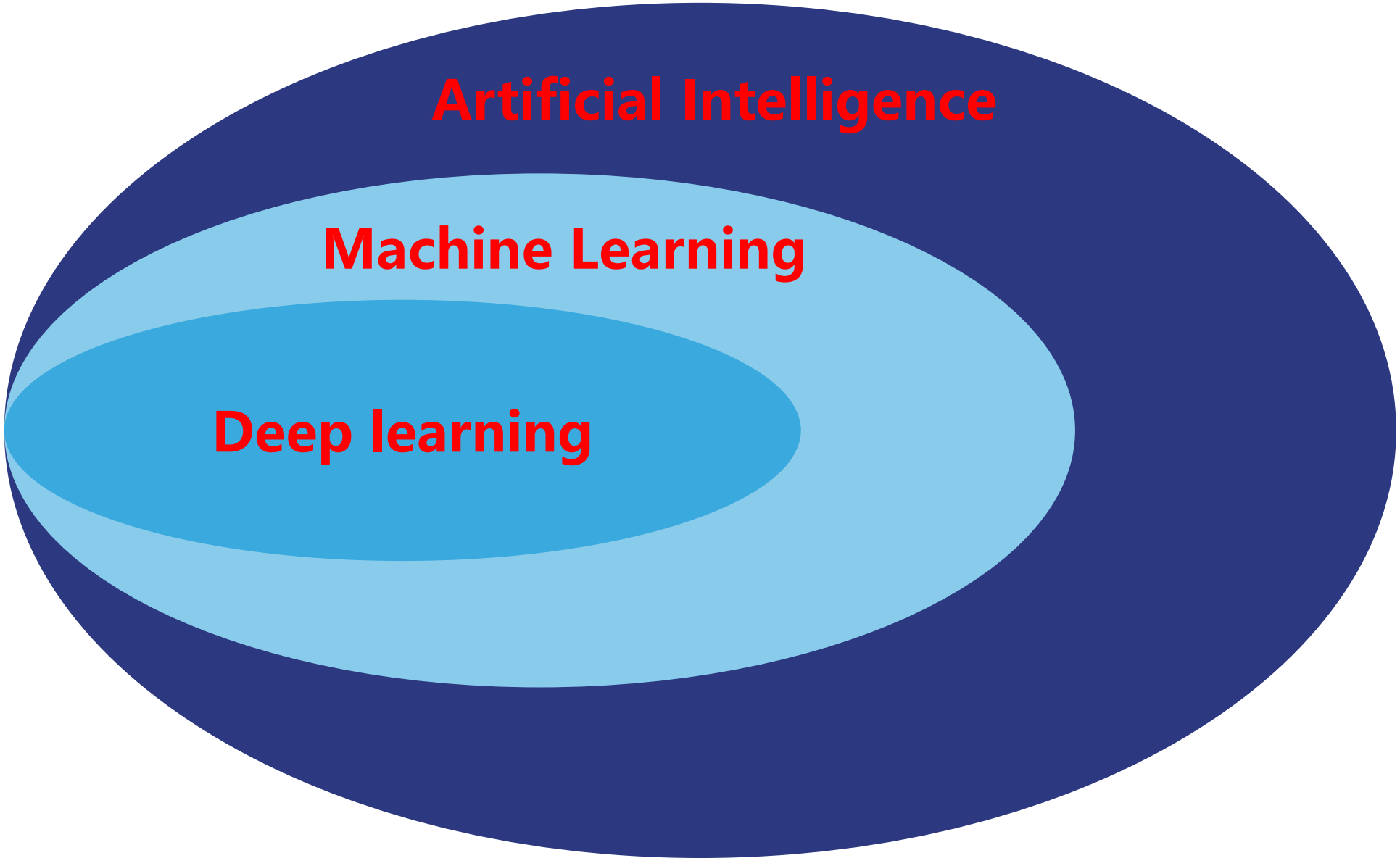
Machine Learning

Machine learning is a model to realize AI.

Deep Learning

DL is an important way of machine learning.

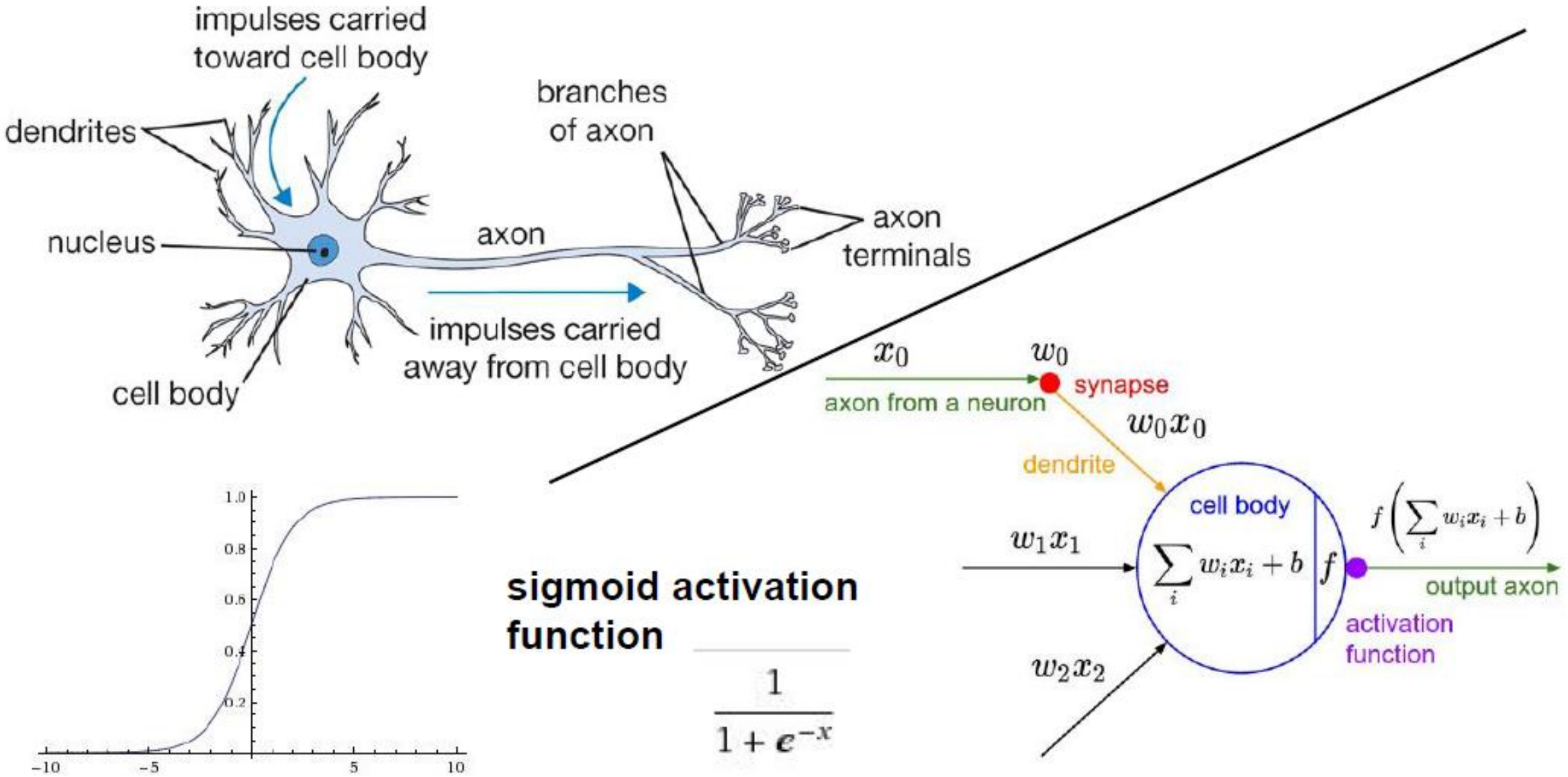
AI,Machine Learning AND Deep Learning



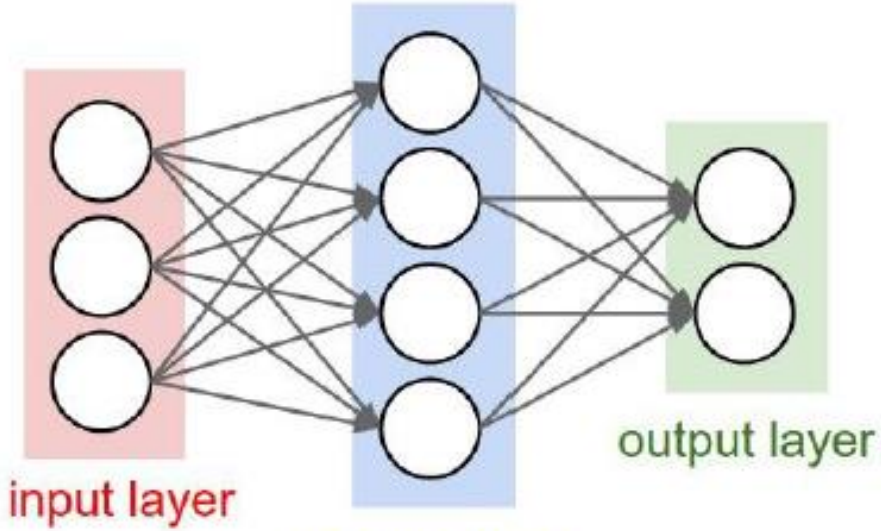
02

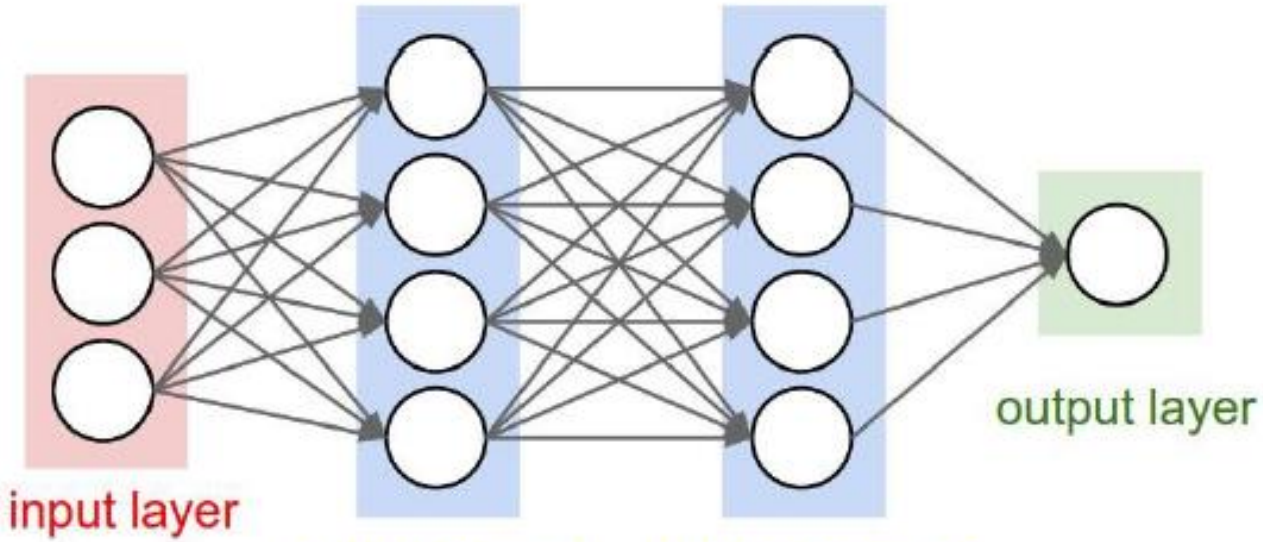
A brief introduction of deep learning

The basic element of neural network



Architectures of neural networks



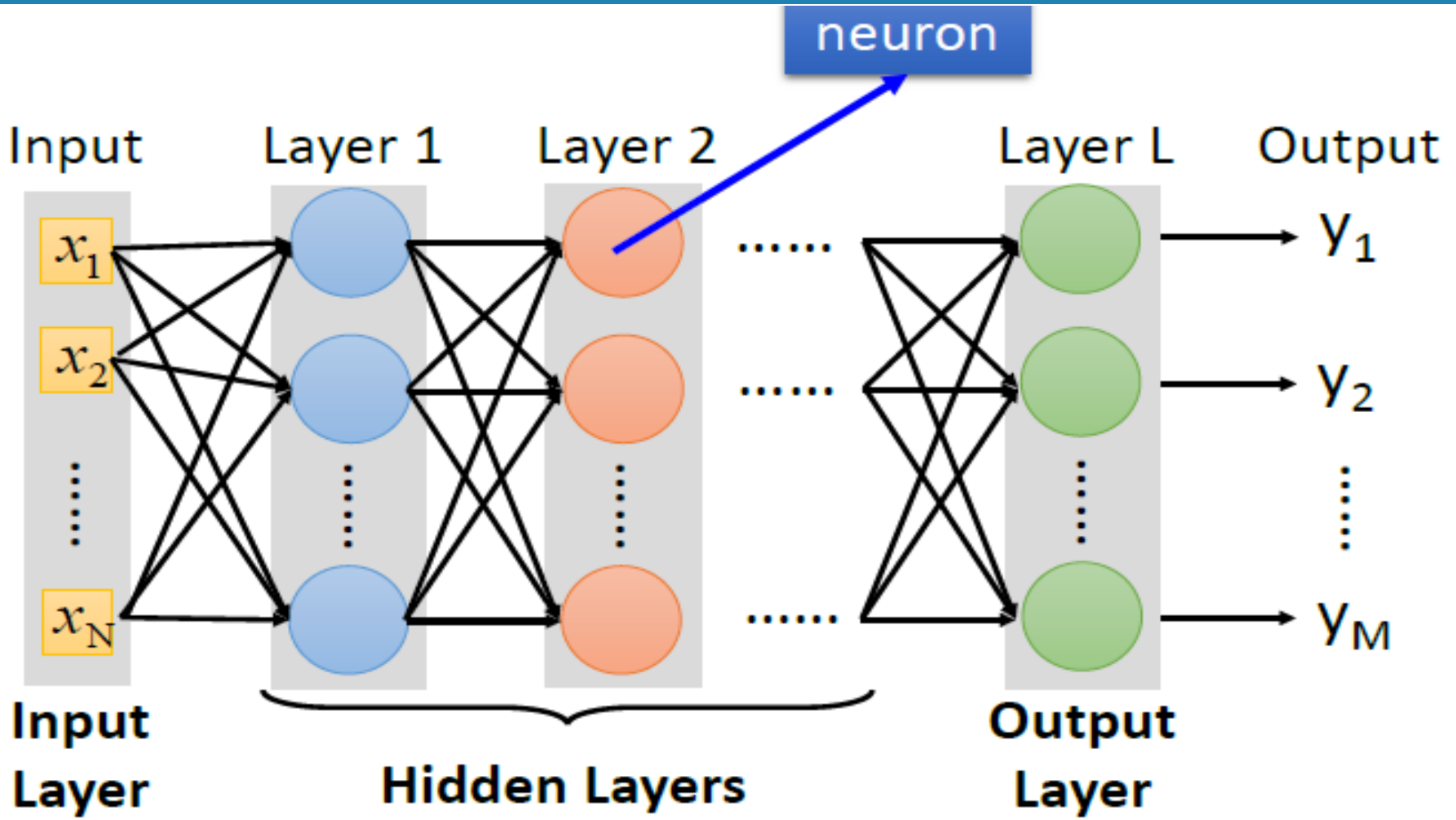


"2-layer Neural Net", or
"1-hidden-layer Neural Net"

"3-layer Neural Net", or
"2-hidden-layer Neural Net"

"Fully-connected" layers

Deep neural network(DNN):Many hidden layers.



Deep means many hidden layers

03

Apply DNN to the flavor tag work

Flavor tag(VTX_08mm):

01 Data set

200k training events

50k test events

02 Input parameters

We totally have 68 parameters.

'aux','vtxlen1','vtxlen1_j','z0cprob','d0bprob' are removed.

And then we have 63 inputs.

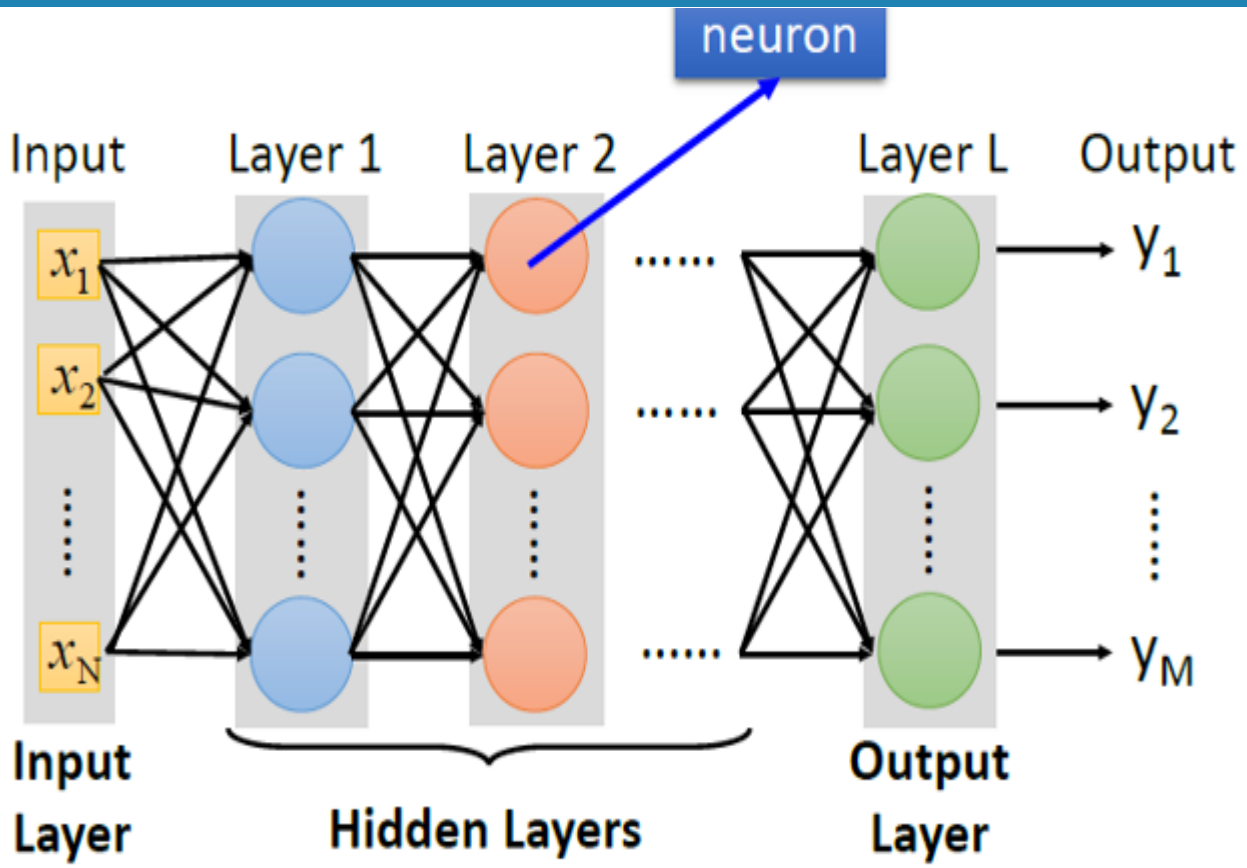
03 Output

Three outputs represent b,c and other tags.

Total parameters

```
['aux' 'ntrkwitho' 'ntrk' 'nvtxall' 'vtxmassal' 'vtxlen12a' 'vtxlen12a'
 '1vtxprob' 'nvtx' 'jete' 'jetcosine' 'jeteta' 'vtxlen1' 'vtxlen2'
 'vtxlen12' 'vtxlen1_j' 'vtxlen2_j' 'vtxlen12_' 'vtxsig1' 'vtxsig2'
 'vtxsig12' 'vtxsig1_j' 'vtxsig2_j' 'vtxsig12_' 'vtxdirang' 'vtxdirang'
 'vtxdirang' 'vtxdirang' 'vtxdirang' 'vtxdirang' 'vtxmom' 'vtxmom1'
 'vtxmom2' 'vtxmom_je' 'vtxmom1_j' 'vtxmom2_j' 'vtxmass' 'vtxmass1'
 'vtxmass2' 'vtxmasspc' 'vtxmult' 'vtxmult1' 'vtxmult2' 'vtxprob'
 'trk1d0sig' 'trk2d0sig' 'trk1z0sig' 'trk2z0sig' 'trk1pt' 'trk2pt'
 'trk1pt_je' 'trk2pt_je' 'jprobr' 'jprobz' 'jprobr5si' 'jprobz5si'
 'sphericit' 'Fd' 'jetrho' 'trkmass' 'nmuon' 'nelectron' 'd0bprob'
 'd0cprob' 'd0qprob' 'z0bprob' 'z0cprob' 'z0qprob']
```

DNN structure for flavor tag



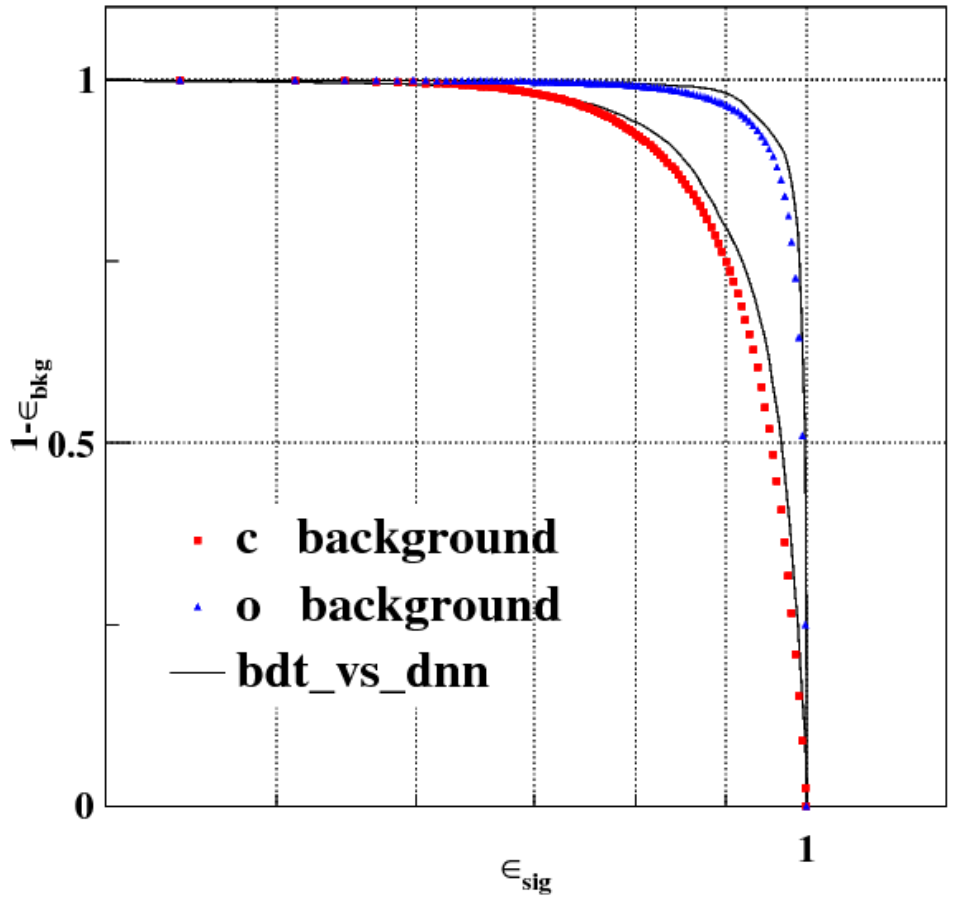
Layer 1(Input layer):63 neurons
Layer 2:512 neurons dropout:0.5
Layer 3:512 neurons dropout:0.5
Layer 4:512 neurons dropout:0.5
Layer 5:512 neurons dropout:0.5
Layer 6:256 neurons dropout:0.5
Layer 7:256 neurons dropout:0.5
Output layer:3 neurons

dropout=0.5 means that the probability that neurons are removed during training is 0.5

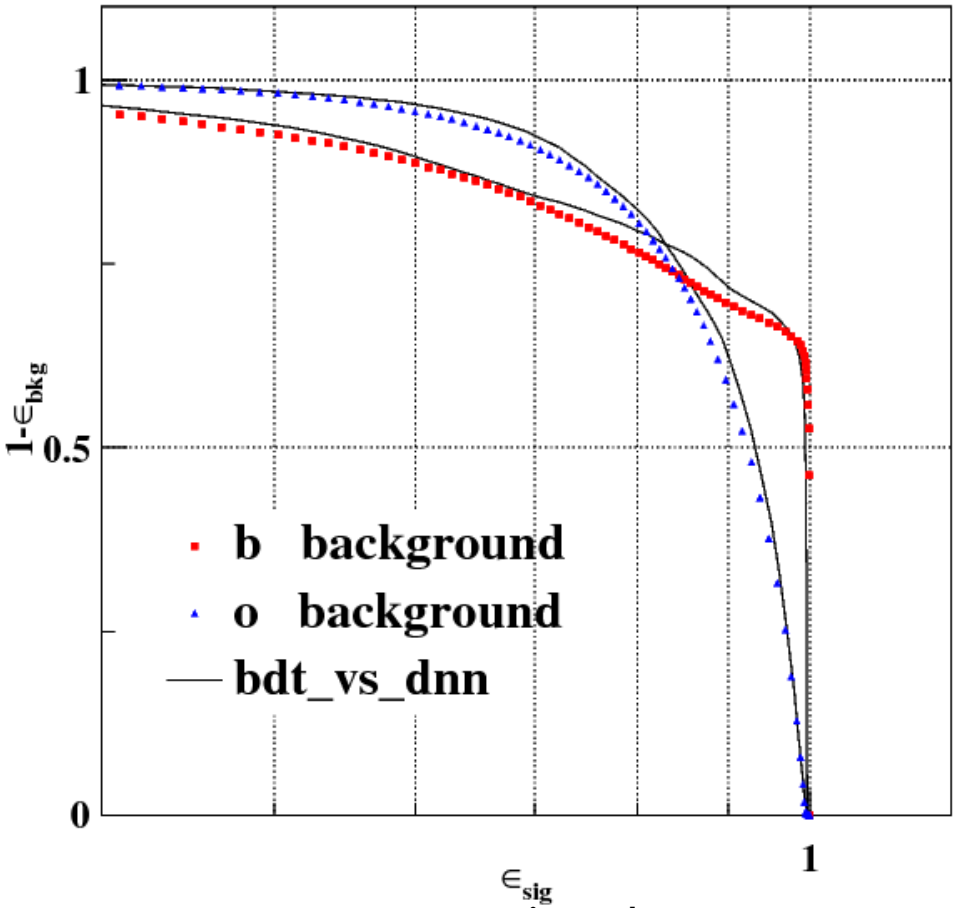
- The entire dataset of 250000 was normalized prior to any training.
- Network parameters were trained using gradient descent with mini-batches of 100 examples per step and 300000 steps.
- The neural networks were trained using the tensorflow package.

ROC compared with bdt:

BDT:630k events DNN:250k events



b signal



c signal

Signal Efficiency & Background Rejection

b signal

E_{sig}(b)	1-E_{bkg}(c, dnn)	1-E_{bkg}(c, bdt)	1-E_{bkg}(o, dnn)	1-E_{bkg}(o, bdt)
0.8	0.94	0.92-0.93	0.99	0.99
0.9	0.79-0.80	0.75-0.76	0.98	0.96-0.97
0.95	0.62-0.65	0.52-0.55	0.93-0.94	0.90-0.91

c signal

E_{sig}(c)	1-E_{bkg}(b, dnn)	1-E_{bkg}(b, bdt)	1-E_{bkg}(o, dnn)	1-E_{bkg}(o, bdt)
0.8	0.8	0.77	0.83	0.81
0.9	0.72	0.7	0.62	0.59
0.95	0.68	0.67	0.4	0.38

Conclusion

For flavor tagging,DNN can improve the background rejection compared with BDT in the same signal efficiency.

Deep learning may be a promising method to deal with other classification problems with lots of physical parameters.