P1:

1.什么是机器学习：

（1）找一个函数：

①speech recognition：一段录音-》“how are you”

②image recognition：一张图片-》类别“cat”

③playing Go：一张棋图-》“5-5”（next move）

2.Deep Learning： f为一个类神经网络

|  |  |  |
| --- | --- | --- |
| 输入-》 | f | -》输出 |
| Vector，  matrix，（image）  sequence（speech，text） |  | Scalar（数字）-》regression  类别-》classification  Text，image |

3.教机器的种种方法：

（1）supervised learning（宝可梦lecture1-5）

需要大量标记资料

（2）self-supervised learning

Pre-train（使用unlabled data，eg：BERT）

Fine-tune

Downstream tasks

（3）generative adversarial network（lecture 6）

x-》function-》y（其中x,y是unpaired）

（4）reinforcement learning（lecture 12）

当不知道怎么label the data，但知道result的好坏时（eg.下围棋）

（5）anomaly detection（异常检测，lecture 8）

让机器具备回答我不知道的能力

（6）explainable AI（lecture 9）

不仅让机器回答类别是宝可梦，也让机器回答为什么类别是宝可梦

（7）model attack（lecture 10）

添加杂讯

（8）domain adaption（lecture 11）

（9）network Compression（lecture 13）

（10）life-long learning（lecture 14）

（11）meta-learning（few-shot learning学习如何学习 lecture 15）

P2 homework 要求

P3 Machine learning≈looking for functions

1. function的types：

（1）regression

（2）classification

（3）structured learning 产生有结构的文件（eg.image,document）

2．怎么找function，三个步骤

（1）写出带有未知参数的公式：

（2）定义Loss：L（w，b）

MAE,MSE,Cross-Entropy

（3）optimization：

Gradient descent（真正的痛点不是local min？？？）

P4 运用activate function：（soft ）sigmoid来找function的步骤

1.未知数的函式

（1）Sigmoid 函式：

（2） 只有一个feature ：

（3） 多个feature

（2）（3）用多个sigmoid函式来逼近piecewise linear curves（分段线性函数）

（3）式可由矩阵表达式计算： 其中b，c，W，x均为矩阵。

2．loss：L（θ）

，其中e是与之间的差距。

1. optimization：

gradient descent 更新θ

* update 和epoch的差别：更新一次θ就是一次update，一般用了一个batch。

所有batch遍历完了之后就是一次epoch。

* ReLU：



2i：两个ReLU叠起来就可以得到一个sigmoid

* 其中sigmoid/ReLU :Neuron/hidden layer

P6 pytorch

1．Training，validation，testing

2．Dataset：stores data

Dataloader：groups data in batches

（1）

（2）

读数据，

返回数据中index位sample

返回数据的size

3．Tensor

(1).shape()

|  |  |  |
| --- | --- | --- |
|  |  |  |
| (5, ) | (3,5) | (高，长，宽) |
| Dim0 | Dim0，dim1 | Dim0，dim1，dim2 |

（2）一些常用：





（3）common operation

①addition：

②subtraction：

③power：

④summation：

⑤mean：

⑥transpose： 将dim0与dim1的shape数互换

⑦squeeze： 将dim0的shape数除去

 在dim1处插入shape=1

⑧cat：在dim1连接三个tensor

（4）DataType：float…

（5）Device：在cpu或gpu上运行



 ------------check GPU

4．Gradient calculation



5．Network layer

（1）Linear Layer(全连接层)



参数：weight,bias

（2）non-linear activation function

Sigmoid Activation:

ReLU Activation: 

6．Build neural network



7．Loss function

For regression：

For classification：



8．Optimization



对于每一个batch：

（1）call  来给模型参数梯度归零

（2）call  回推算每一层的梯度

（3）call  根据 （2）算出来的梯度调整模型参数

9．完整步骤

（1）training setup



（2）training loop



（3）validation loop



（4）testing loop



保存模型：

加载模型：

P7 pytorch 实验以及易犯的errors

P9同P6

P10同P7

P11 homework1讲解

主要调net结构和config参数

P13 deep learning 简介

1. Deep learning 的历史（ups and downs）
2. Three steps，similar to the machine learning
3. neuron连接方式：
   1. fully collected 全连接
   2. CNN
4. Matrix operation 矩阵运算GPU加速
5. Layer的个数及其neuron的个数由自己调
6. 为什么deep而不是fat？？？

P 14 back propagation

1. Gradient descent

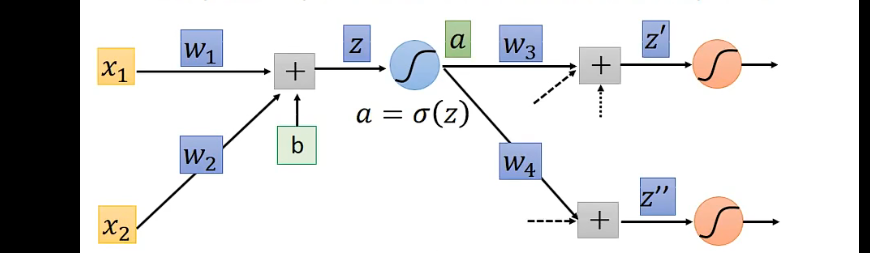


1. Backpropagation，to compute the gradient efficiently
2. Chain rule：

Case1: Case2:

4．。

一层Neural network：





（1）

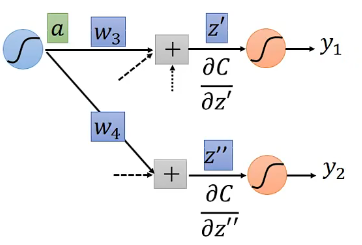
（2）根据chain rule：，其中为forward pass，为backward pass。

（3）

（4） 其中，

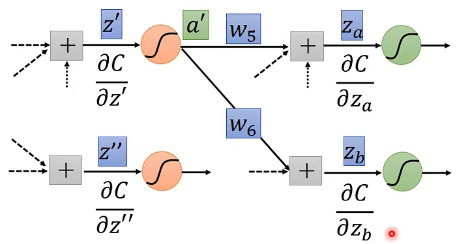
（5）则

Case1:后面直接是output layer





Case2：not output layer



因为（5），只要知道就可以计算出。同理可有计算得出，如此重复直到出现case1的情况为止。如下：

