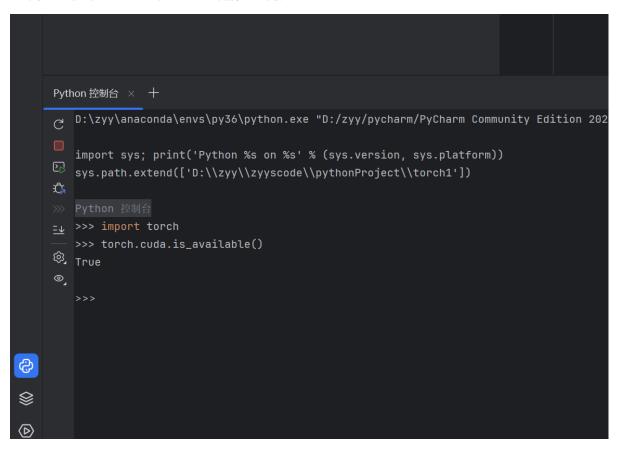
配置了一个显卡驱动pytorch的环境,但是遇到了很多问题。

在anaconda配置python=3.6的pytorch环境再配置到pycharm解释器之后import torch仍然报错。折腾了两天后发现自己查看cuda版本的指令用的是nvcc --version,返回的是12.1,然而这是cuda编译器的版本,更改指令nvidia-smi之后发现自己其实是12.7版本,而pytorch目前还不能下载到适配12.7cuda的版本,重新安装了12.1版本CUDA之后解决了问题

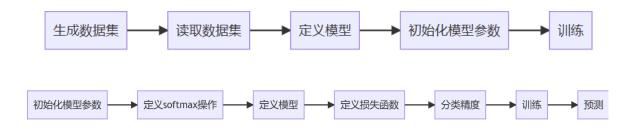


通过调包可以简洁实现许多模型 不要重复发明轮子

• MNIST数据集

是图像分类中广泛使用的数据集之一,但作为基准数据集过于简单。 课程使用类似但更复杂的 Fashion-MNIST数据集。

正常模型训练流程



多层感知机

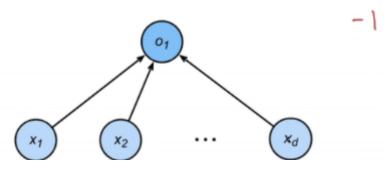
感知机

感知机是一个二分类模型

虽然单层感知机有其局限性,但多层感知机 (MLP) 通过引入隐藏层和非线性激活函数,可以解决更复杂的任务

·给定输入x,权重w,和偏移b,感知机输出:

$$o = \sigma \left(\langle \mathbf{w}, \mathbf{x} \rangle + b \right) \quad \sigma(x) = \begin{cases} 1 & \text{if } x > 0 \\ \mathbf{0} & \text{otherwise} \end{cases}$$



- **感知机与回归和Softmax回归的区别** 感知机是二分类(1或-1),而回归的输出是实数,Softmax回归的输出是概率。
- 训练感知机

initialize
$$w = 0$$
 and $b = 0$
repeat
if $y_i [\langle w, x_i \rangle + b] \leq 0$ then
 $w \leftarrow w + y_i x_i$ and $b \leftarrow b + y_i$
end if
until all classified correctly

等价于使用批量大小为1的梯度下降, 并使用如下的损失函数

$$\ell(y, \mathbf{x}, \mathbf{w}) = \max(0, -y\langle \mathbf{w}, \mathbf{x} \rangle)$$

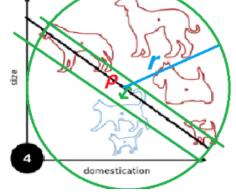
收敛定理

- ・数据在半径r内
- · 余量 ρ 分类两类

$$y(\mathbf{x}^{\mathsf{T}}\mathbf{w} + b) \ge \rho$$

对于
$$\|\mathbf{w}\|^2 + b^2 \le 1$$

. 感知机保证在 $\frac{r^2 + 1}{a^2}$ 步后收敛

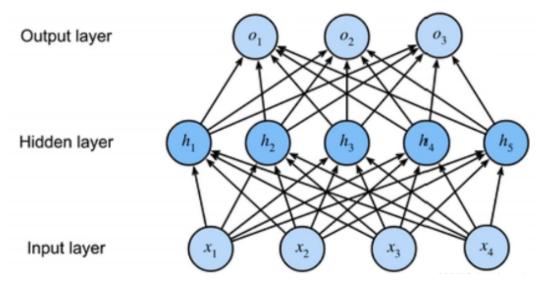


多层感知机

类似于softmax函数添加隐藏层

softmax函数类似于输出层,将数据输入转化为预测为各个分类的概率,感知机解决二分类问题,多层感知机则解决了XOR问题

隐藏层大小是超参数 (输入和输出大小由数据决定,输出大小人为决定。)



激活函数

①Sigmoid 激活函数**②Tanh 激活函数③ReLU 激活函数

总结

- 多层感知机使用隐藏层和激活函数来得到非线性模型
- ·常用激活函数是Sigmoid,Tanh,ReLU
- · 使用 Softmax 来处理多类分类
- · 超参数为隐藏层数,和各个隐藏层大小 Michanyia owai

最后在试图安装d2l库复现多层感知机的代码时pip install失效,上网找了切换国内源的链接之后仍然无法安装,最后电脑在conda update -all途中卡顿,退出重进之后整个环境就崩溃了

目前还没找到解决办法,所以最后代码也没打成,再不行就删掉重配了

之前那个比赛我分到做ppt的任务花去一部分时间,配置环境又消磨了大部分的时间,所以本周学习内容较少