

## 神经网络课程作业 3

1. 什么是单层感知机的局限性，造成其局限性的原因是什么，该如何解决？
2. 相比单层感知机，多层感知机有什么优点，又带来什么问题？
3. 比较绝对误差与平方误差的优劣，并介绍其他的任意两种误差函数。
4. 描述梯度下降的优缺点和局部最小值的定义，使用梯度下降方法优化 Himmelblau 函数： $f(x,y) = (x^2 + y - 11)^2 + (x + y^2 - 7)^2$ 。可使用 Pytorch 框架。（Himmelblau 函数可视化代码如下）

```
import numpy as np
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import Axes3D
def him(x):
    return (x[0]**2+x[1]-11)**2 + (x[0]+x[1]**2-7)**2

x = np.arange(-6,6,0.1)
y = np.arange(-6,6,0.1)
X,Y = np.meshgrid(x,y)
Z = him([X,Y])

fig = plt.figure()
ax = fig.gca(projection='3d')
ax.plot_surface(X,Y,Z,cmap='rainbow')
ax.set_xlabel('x[0]')
ax.set_ylabel('x[1]')
ax.set_zlabel('f')
fig.show()
```

5. 代码设计单层感知机，使其满足  $y = \sigma(Wx + b)$ 。（其函数定义如以下代码，只需完成其实现）

```

class Layer():
    def __init__(self,input_dim,output_dim,bias=True):
        None
    def __call__(self,x,train=False):
        None
    def forward(self,x,train):
        None
    #反向传播函数
    def backward(self,error,eta):
        None

```

6. 在第 5 题的基础上，使用 Layer 类完成多层感知机的构造，其成员函数和 Layer 一致，并用于对  $\cos(2\pi x)$  函数的拟合。

注意：对于 4、5、6 题，请在 pdf 文件中贴代码，并配以必要的注释。