X2

X3

X4

H2

H3

H1

Y2

Y1

参数共享模型创新简易证明：

E2

E1

W13

W1

W3

W2

W4

X1

对于整体损失函数：



反向传播过程中，对模型参数w1求梯度如下：



若 >0，且<0

当 >0时

更新 w1 = w1-\*步长

此时，对于任务Y1，损失函数E1 减小；对于任务Y2，损失函数E2增加；整体损失函数E减小。

反之，当<0时，

此时，对于任务Y1，损失函数E1 增加；对于任务Y2，损失函数E2减小；整体损失函数E减小。

我们把在多任务学习的反向传播过程中，共享层权重参数对于不同任务的损失函数的梯度出现异号的情况称之为梯度异号。

通过求解共享层权重对于各任务损失函数的梯度表达式，易证，对于极大极小值归一化处理后的数据，若w1出现梯度异号的情况，则同样在共享曾连接H1结点的w2,w3,w4也存在梯度异号的情况。

此时将原本模型调整如下：

图示

描述已自动生成

令 ,的初值依旧分别等于w1,w2,w3,w4。

此时H1为只针对Y1任务的单任务节点，H1 为只为Y2任务的单任务节点。

对于 w1 = w1-\*步长

对于=-\*步长

此时，不论大于0或小于0，对于任务Y1，若损失函数E1对于w1的梯度在一定范围内单调不变，则损失函数E1 更大幅度减小；对于任务Y2，损失函数E2依旧减小；整体损失函数E更大幅度减小。

图表, 雷达图

描述已自动生成

图表, 折线图

描述已自动生成

算法简易实现：

1、设置迭代次数阈值 ，公共结点为H={H1,H2,H3,H4}

2、对于原始模型进行迭代训练，计算各个子任务损失函数对权重w的梯度方向，若多个子任务损失函数对于权重w的梯度方向为异号且方向连续不变，经过t轮迭代依旧如此，若t>，复制w对应的公共节点为H”，H“与父节点之间依旧全连接，其加法器的权重w初始值不变，H”单独只与其中一个子任务连接，被复制的公共节点H也只与另一个单独的子任务连接。

3、对于更新后的网络继续进行迭代训练，达到最优解。

