



推理中的丢弃法

- 正则项只在训练中使用：他们影响模型参数的更新
- 在推理过程中，丢弃法直接返回输入

$$\mathbf{h} = \text{dropout}(\mathbf{h})$$

- 这样也能保证确定性的输出

总结



- 丢弃法将一些输出项随机置0来控制模型复杂度
- 常作用在多层感知机的隐藏层输出上
- 丢弃概率是控制模型复杂度的超参数



nn.ReLU()之后