

Variance 刻画扰动造成的影响

Bias 刻画学习算法本身的拟合能力

Noise 当前任务任何学习算法所能达到的误差下界

L1 Regularisation Term $\sum_{i=1}^n |w_i|$

L2 Regularization Term $\sum_{i=1}^n w_i^2$

↑
Weight decay 权重衰减

Dropout: 随机失活

`torch.nn.Dropout()`

p : 被舍弃的概率