

其他训练Tricks

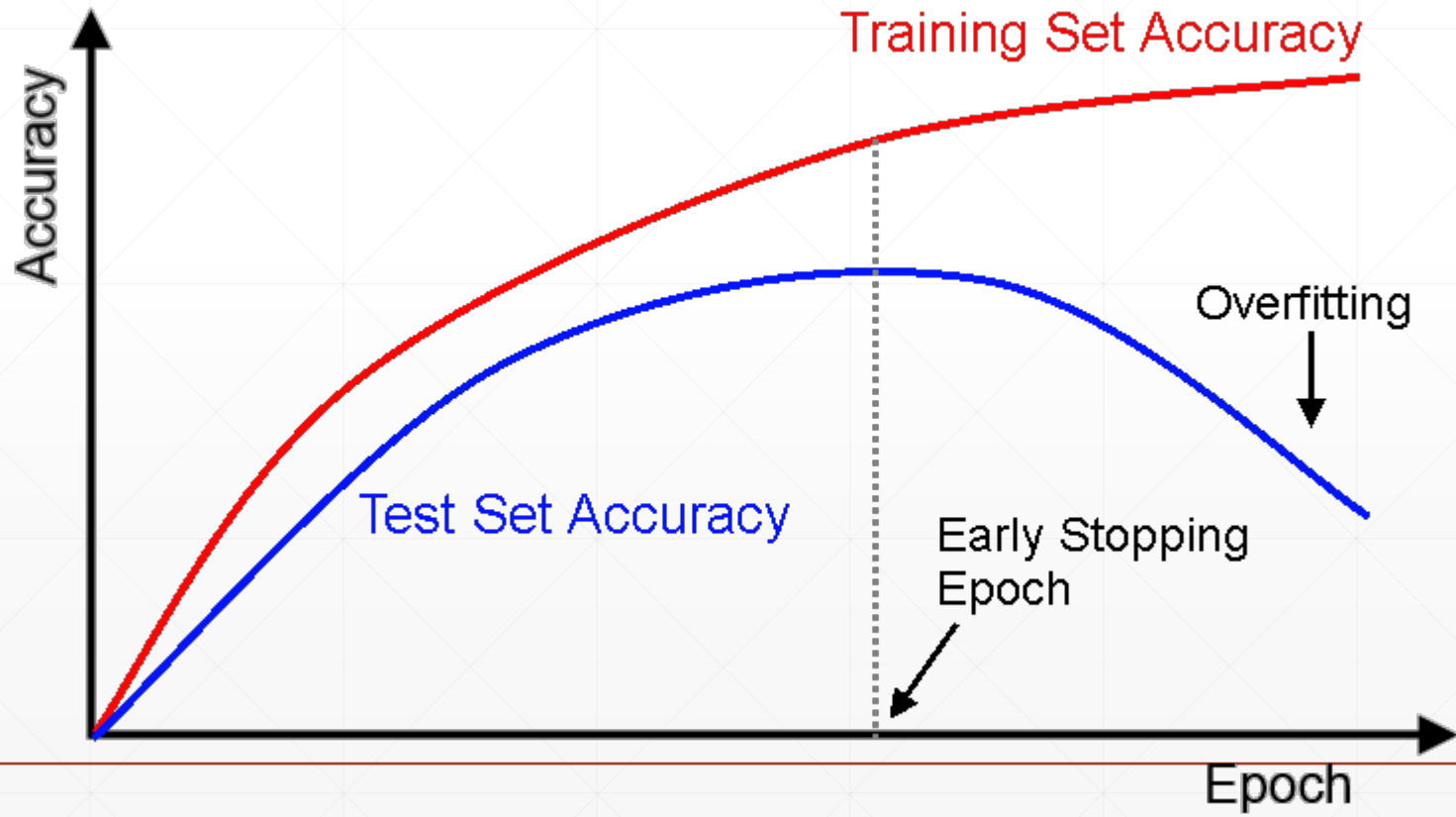
主讲：龙良曲

Tricks

- Early Stopping
 - Dropout
 - Stochastic Gradient Descent
-

Early Stopping

- Regularization

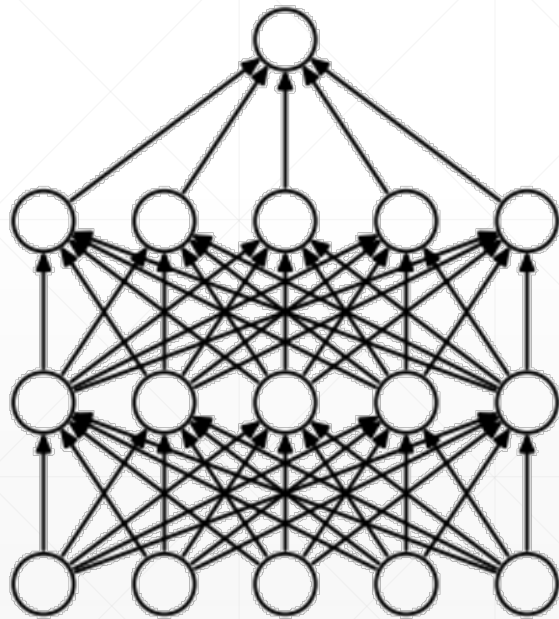


How-To

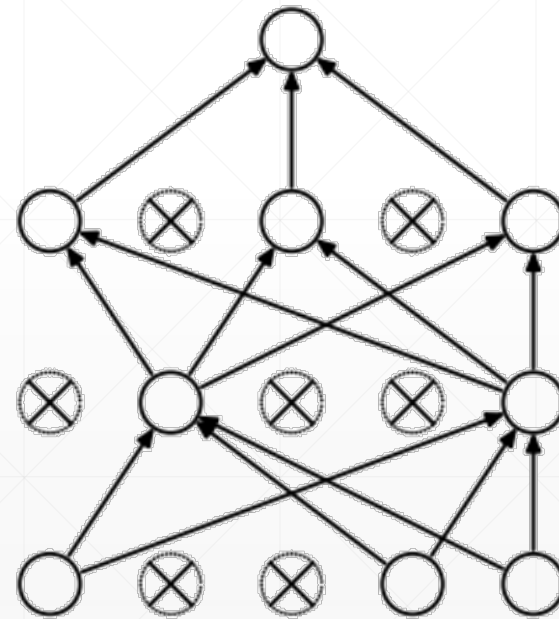
- Validation set to select parameters
 - Monitor validation performance
 - Stop at the highest val perf.
-

Dropout

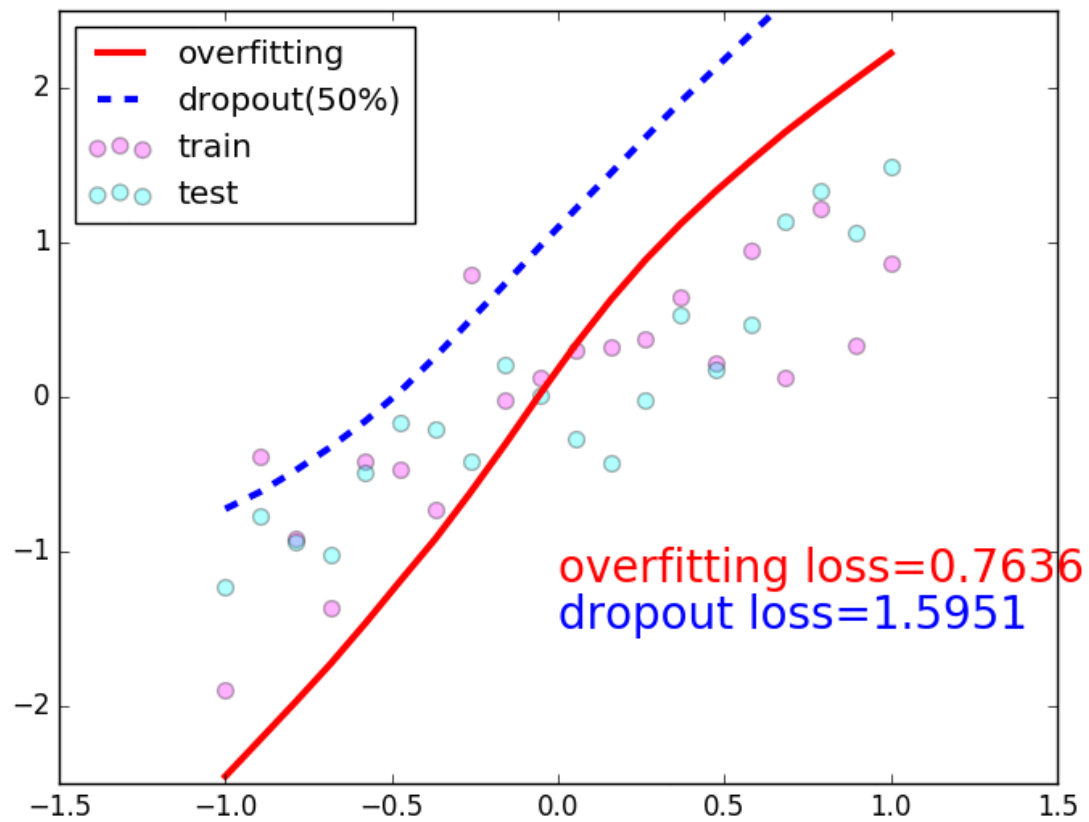
- *Learning less to learn better*
- Each connection has $p = [0, 1]$ to lose



(a) Standard Neural Net



(b) After applying dropout.





```
network = Sequential([layers.Dense(256, activation='relu'),  
                      layers.Dropout(0.5), # 0.5 rate to drop  
                      layers.Dense(128, activation='relu'),  
                      layers.Dropout(0.5), # 0.5 rate to drop  
                      layers.Dense(64, activation='relu'),  
                      layers.Dense(32, activation='relu'),  
                      layers.Dense(10)])
```

API

- `layer.Dropout(rate)`
 - `tf.nn.dropout(x, rate)`
-

Behavior between train and test



Batch-
Norm

```
● ● ●  
for step, (x,y) in enumerate(db):  
  
    with tf.GradientTape() as tape:  
        # [b, 28, 28] => [b, 784]  
        x = tf.reshape(x, (-1, 28*28))  
        # [b, 784] => [b, 10]  
        out = network(x, training=True)  
  
    # test  
    out = network(x, training=False)
```

**JUST
DO
IT.**

Stochastic Gradient Descent

- Stochastic
 - not random!
- Deterministic



Gradient Descent

$$\frac{\partial}{\partial \theta_j} J(\theta) = \frac{1}{m} \sum_{i=1}^m (\hat{y}^i - y^i) \cdot x_j^i$$

② Vanilla (Batch) G.D.

$$\theta_j := \theta_j - \alpha \cdot \underbrace{\frac{\partial}{\partial \theta_j} J(\theta)}_{\frac{1}{m} \sum_{i=1}^m (\hat{y}^i - y^i) x_j^i}$$

Gradient Descent

③ Stochastic G.D.

for i in range (M):

$$\theta_j := \theta_j - \alpha \cdot \frac{\nabla J}{\nabla \theta_j} (\hat{y}^i - y^i) x_j^i$$

only one example

Stochastic Gradient Descent

- Not single usually
 - batch = 16, 32, 64, 128...
-

Why

索泰
ZOTAC



¥1789.00

索泰 (ZOTAC) GeForce GTX1060 X-GAMING OC吃鸡显卡/游戏电竞台式机独

6.6万+条评价

二手有售

索泰京东自营旗舰店

自营 秒杀 券1500-200

索泰
ZOTAC



¥7599.00

索泰 (ZOTAC) GTX1080Ti-11GD5X至尊PLUS OC绝地求生/吃鸡显卡

40+条评价

云轩 DIY整机专营店

满减 免邮 满199-100 险

NVIDIA
TESLA

NVIDIA
CUDA



¥80999.00

NVIDIA TESLA V100 16GB/32GB HBM2/
人工智能深度学习GPU运算 16GB HBM2

0条评价

子希DIY电脑专营店

免邮

下一课时

卷积神经网络

Thank You.
